

Catalytic coherence

Johan Åberg*

Institute for Physics, University of Freiburg, Hermann-Herder-Strasse 3, D-79104 Freiburg, Germany

Conservation of energy implies that we cannot freely create superpositions between energy eigenstates, while access to coherence enables such processes. A simple example of the latter is when an atom is put into a superposition of two energy levels, using coherent radiation in form of a laser. The question is to what extent the coherence is degraded when utilized. Here we show that coherence can be turned into a catalytic resource, meaning that it can be used indefinitely to induce coherent operations without ever being degraded. As an application we consider quantum thermodynamics, where we demonstrate that access to sufficient coherence enables the release of all latent expected work content in a system. Since we can make it a catalyst, no coherence is spent to achieve this work extraction. In a wider perspective, the fact that coherence can be made catalytic provides a step towards a characterization of the relation between energy and coherence as resources in quantum statistical mechanics.

Introduction.— Coherence is a fundamental resource that enables us to implement coherent operations on quantum systems; a canonical example being the use of a laser to put an atom in superposition between two energy-levels, or radio-pulses for nuclear spins. Due to conservation of energy, we cannot directly turn a quantum system with a definite energy into a superposition of different energies. However, with access to a resource in form of a system with a high degree of coherence we can overcome this limitation to an arbitrarily good approximation. The question is to what extent this ‘coherence resource’ gets degraded when utilized.

Here we show that coherence can be turned into a catalyst, meaning that it can be used repeatedly without ever being degraded. We consider two models with this catalytic property. The first of these models is convenient to analyze, but is somewhat unphysical in that it has no ground state. The second model amends this problem, at the price of a protocol that regularly injects energy into the system. To touch upon the question of the existence of these effects in more realistic systems, we furthermore make a brief numerical analysis of the Jaynes-Cummings model [1, 2].

As an application we turn to the question of ‘work extraction’, i.e. how much work that can be extracted from a system in a non-equilibrium state with respect to a heat bath of given temperature. The work extraction problem, and the closely related concepts of information erasure and Maxwell’s demon, have a long history (see e.g., [3–8]) with a recent revived interests, e.g., in the contexts of resource theories [9–11] and single-shot statistical mechanics [12–18]. Recently, it was demonstrated [19] that for systems with non-classical states, meaning superpositions between different energies, not all the expected work content can be extracted, if we are restricted to operate on the systems individually, and have no access to additional coherence resources. Here we show that access to sufficient coherence does enable the release of all the

latent expected work content of a system, even for individual operations. Since this analysis is performed within a model where coherence is catalytic, this implies that the additional release of expected work is not paid for by a loss of coherence. Finally, we make some observations in relation to single-shot work-extraction [12–18], which raise questions concerning the relation between coherence and noise-free energy. In relation to this thermodynamic application, one may note [20], which used a coherent extraction device as a negentropy source, to demonstrate a transient efficiency that exceeds the standard Carnot bound for work extraction against a hot and a cold heat bath.

Coherence as a resource.— Imagine a two-level system S with energy eigenstates $|\psi_0\rangle$ and $|\psi_1\rangle$, with corresponding energy eigenvalues h_0 and h_1 with $h_1 > h_0$. If we wish to perform non-trivial operations on S , we need, due to conservation of energy, another system where we can withdraw and deposit energy. We refer to this as the ‘energy reservoir’, E , with Hamiltonian H_E . If the spectra of H_S and H_E are suitably matched, i.e., if there are resonances, then the combined Hamiltonian $H_S + H_E$ have degeneracies. (For the sake of notational simplicity we avoid to write the more correct expression $H_S \otimes \hat{1}_E + \hat{1}_S \otimes H_E$.) This means that the ground state $|\psi_0\rangle$ can be excited into $|\psi_1\rangle$ if we simultaneously de-excite the energy reservoir. But what if we wish to transform the the ground state $|\psi_0\rangle$ into a superposition $(|\psi_0\rangle + |\psi_1\rangle)/\sqrt{2}$? With a bit of thought one can realize that it is impossible to do this with an energy preserving unitary operation if both the system and the reservoir initially are in an energy eigenstate. (This observation can be understood in the much wider context of ‘reference frames’ and symmetry preserving operations [21–23].) At first sight this might seem a bit odd, since putting systems into superpositions between energy eigenstates is a common practice in many labs. As is well known, one way to resolve this apparent paradox is to realize that these operations usually are induced by laser or radio-fields that can be modeled as coherent states, which are broad superpositions of the energy eigenstates (the number states) of a mode of the electromagnetic field [24–26] (although this

*Electronic address: johan.aberg@physik.uni-freiburg.de

can be debated, see e.g. [27–32]). Since coherence thus is a resource that allows us to perform non-trivial quantum operations, it appears a relevant question to ask to what extent it gets degraded when we use it. In the following we shall consider a model where it can be shown that the coherence does not degrade at all. In other words, it is a catalyst.

The doubly-infinite energy-ladder.– The model we use has previously been considered in the context of quantum thermodynamics, for the purpose of work extraction [19]. The Hamiltonian of the reservoir in this model has the form

$$H_E^{(s)} = s \sum_{j \in \mathbb{Z}} j |j\rangle\langle j|, \quad (1)$$

where $s > 0$ is the energy spacing in this doubly infinite ladder, and $\{|j\rangle\}_{j \in \mathbb{Z}}$ is an orthonormal basis, where \mathbb{Z} denotes the set of (positive and negative) integers. Regarded as a Hamiltonian, $H_E^{(s)}$ is slightly odd, in that it does not have any ground state. We shall shortly remedy this problem, but due to several convenient properties of this model (see Appendix A 1) we shall use it for the initial analysis. We let S be a two-level system with $h_1 = s$ and $h_0 = 0$, and thus S and E are in resonance. (See Appendix A for the case of N -level systems.) Irrespective of which state the energy reservoir is in, it can always compensate for a transition on S . Next, we define the unitary operator $\Delta = \sum_{j \in \mathbb{Z}} |j+1\rangle\langle j|$. As seen, this operator translates the state ‘rigidly’ along the energy ladder. With the help of Δ we can formulate the following family of unitary operators on $\mathcal{H}_S \otimes \mathcal{H}_E$

$$\begin{aligned} V(U) &= \sum_{n, n'=0,1} |\psi_n\rangle\langle\psi_n| U |\psi_{n'}\rangle\langle\psi_{n'}| \otimes \Delta^{n'-n} \\ &= \sum_{j \in \mathbb{Z}} V_j(U), \\ V_j(U) &= \sum_{n, n'=0,1} |\psi_n\rangle\langle\psi_n| Q |\psi_{n'}\rangle\langle\psi_{n'}| \otimes |j-n\rangle\langle j-n'|, \end{aligned} \quad (2)$$

where U is an arbitrary unitary operator on \mathcal{H}_S . As one can see, the mapping V does in some sense create an infinite number of copies of U on the combined space $\mathcal{H}_S \otimes \mathcal{H}_E$. By construction, all $V(U)$ commute with $H_S + H_E$, i.e., they are energy conserving (see Appendix A 1). Furthermore they commute with all Δ^a , and thus act uniformly over the energy ladder (see [19] for discussions on this). The family of unitary operators $V(U)$ serves as the set of ‘allowed operations’ in our model. In what follows, we shall investigate what kind of operations that this model allows us to perform on system S , and how this depends on the coherence of the state of the reservoir E . Finally, we show that this set of operations does not change, irrespective of how many times we use the reservoir. In this sense, the coherence of the reservoir is a catalytic resource in this model.

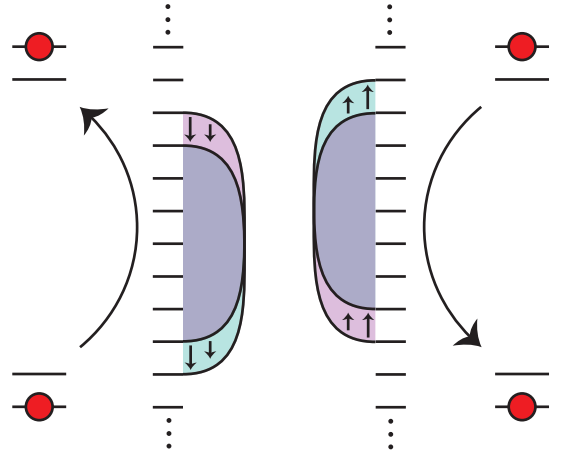


FIG. 1: **Generalized coherent states.** As a simplified versions of coherent states, one can take uniform superpositions over a range of energy levels $|\eta_{L,l_0}\rangle = \sum_{l=0}^{L-1} |l_0+l\rangle/\sqrt{L}$. When interacting with a two-level system, the whole state gets shifted up or down along the ladder, depending on whether the two-level system absorbs or donates energy. For large L , the shifted and unshifted states have an overlap close to 1, i.e., they are almost identical. This makes it possible to approximately create superpositions in the two-level system, and more generally to approximately perform unitary operations. In the limit of large L , these implementations can be made perfect.

Having defined our model, the question is what kind of operations it allows us to perform on S . Given a state σ on the reservoir, we can implement channels on S via

$$\Phi_{\sigma,U}(\rho) = \text{Tr}_E[V(U)\rho \otimes \sigma V(U)^\dagger]. \quad (3)$$

Given a state σ we let $\mathcal{C}(\sigma)$ denote the set of channels $\Phi_{\sigma,U}$ that can be implemented via arbitrary unitary operators U .

Next, we shall demonstrate that if σ is sufficiently coherent, then we can, to a good approximation, implement arbitrary unitary operations on S (see Fig. 1). For this purpose, we consider states $\sigma = |\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|$ of the form

$$|\eta_{L,l_0}\rangle = \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} |l_0+l\rangle, \quad (4)$$

i.e., uniform superpositions over a collection of consecutive energy eigenstates. One can show (Appendix A 3) that in the limit of large L , the channel $\Phi_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|,U}$ converges to the unitary operation $\rho \mapsto U\rho U^\dagger$. In other words, the higher degree of coherence (i.e., the larger L) the better we can implement arbitrary unitary operations on S . Hence, our model is powerful enough to achieve all unitary operations on the system, given a sufficient resource of coherence in the reservoir. One may note that this is analogous to a similar constructions for coherent states of a bosonic mode interacting with an atom (see, e.g., chapter 7 in [33]).

Catalytic coherence in the doubly-infinite ladder.— Let us now take a closer look on what aspects of the state σ it is that determine the set of channels $\mathcal{C}(\sigma)$ that can be reached. If one inserts Eq. (2) into (3), it turns out that $\Phi_{\sigma,U}$ depends on σ only via expectation values of the form $\text{Tr}(\Delta^a \sigma)$ for $a \in \mathbb{Z}$. (See equation (A10) in Appendix A 2.) Hence $\mathcal{C}(\sigma)$ only depends on the set of numbers $\text{Tr}(\Delta^a \sigma)$ (which does not determine σ uniquely).

To investigate the catalytic aspects of the coherence we need to determine how the state of the energy reservoir changes when we use it. For that purpose we define the corresponding channel on E

$$\Lambda_{\rho,U}(\sigma) = \text{Tr}_S[V(U)\rho \otimes \sigma V(U)^\dagger]. \quad (5)$$

By using Eq. (2) one can confirm that

$$\text{Tr}[\Delta^a \Lambda_{\rho,U}(\sigma)] = \text{Tr}(\Delta^a \sigma), \quad (6)$$

for all σ, ρ, U , and a . (See Lemma 5 in Appendix A 4.) In other words, the expectation value of all powers of Δ are invariant under the action of these operations. Above, we noted that the only aspects of σ that determine what channels can be implemented, are precisely these powers $\langle \Delta^a \rangle = \text{Tr}(\Delta^a \sigma)$. We can thus conclude the set of channels that can be implemented the second time we use the reservoir is not affected by the first use, i.e., $\mathcal{C}(\Lambda_{\rho,U}(\sigma)) = \mathcal{C}(\sigma)$. (For the full statement, see Proposition 2 in Appendix A 4.) In other words, we do not degrade the coherence resource in the reservoir by using it. In this sense, coherence is catalytic in this model. One can also prove a stronger catalytic property, which does not assume that the reservoir and the subsequent systems initially are uncorrelated (see Appendix A 5.)

It should be emphasized that although $\langle \Delta^a \rangle$ are invariants, the underlying *state* does change. Typically it appears to get ‘smeared out’ over larger and larger stretches of the energy ladder. (See Appendix A 6 for an example.)

The half-infinite ladder (harmonic oscillator).— Since the state of the reservoir gets broader for each use, one could worry that the catalytic properties of the coherence in the above model is an anomaly, related to the fact that the model lacks a ground state; for a system with ground state the broadening distribution would eventually ‘hit the bottom’. This is indeed true, but here we show that by introducing a very simple protocol, the set of channels that the reservoir can implement, can be maintained indefinitely, also in a model that has a proper ground state. To this end, we ‘cut away’ the lower half of the doubly-infinite ladder and thus obtain (the spectrum of) the harmonic oscillator

$$H_E^+ := s \sum_{j=0}^{+\infty} j|j\rangle\langle j|. \quad (7)$$

We define a new class of unitary operations on S and E as

$$V_+(U) = |\psi_0\rangle\langle\psi_0| \otimes |0\rangle\langle 0| + \sum_{l=1}^{+\infty} V_l(U), \quad (8)$$

with V_l as in equation (2). By direct comparison one can see that $V(U)$ and $V_+(U)$ act in an identical way on all states with at least one quanta, i.e., $\langle l|\sigma|l\rangle = 0$ for $l = 0$. (For multi-level systems this ‘border zone’ would be larger.)

Let us now consider the dynamics in this model. For this purpose, assume that the state σ have a projection only on a finite range of energy levels. Since the two-level system S can absorb or donate at most one energy quanta, it follows that the state on the energy reservoir can extend the support on the energy eigenstates with at most one energy level up as well as down in energy when the systems interact. Imagine now that reservoir initially is in a state with at least k quanta (i.e., $\langle j|\sigma|j\rangle = 0$ for all $j < k$). For k uses of the reservoir on a sequence of two-level systems, the half-infinite ladder would thus be indistinguishable from the doubly-infinite ladder, and thus the set of channels that the reservoir can implement remains constant for these k steps, i.e., a kind of ‘quasi-catalytic’ property. By translating the initial state further and further up in energy (without increasing the ‘width’ of the superposition), we can make the ‘lifetime’ of this quasi-catalytic behaviour arbitrarily long. (See Proposition 4 in Appendix B 2.)

Protocol for a catalytic half-infinite ladder.— Next we shall see that we can achieve better than a mere quasi-catalytic property for the half-infinite ladder model. By a simple protocol we can maintain the coherence properties of the reservoir indefinitely (see Fig. 2 and Appendix B 3 for more details). To this end, we assume an initial state σ_{in} such that $\langle l|\sigma_{\text{in}}|l\rangle = 0$ for $l = 0, 1$, i.e., there are at least two quanta in this state. We let E interact with the first system S_1 with respect to some arbitrary choice of unitary $V_+(U_1)$. As we know from the above reasoning, the effect is identical to $V(U_1)$. We also know that the new state $\bar{\sigma}$ on the energy reservoir is such that $\langle l|\bar{\sigma}|l\rangle = 0$ for $l = 0$. Consider an ancillary two-level system A , with the two eigenstates $|a_1\rangle, |a_0\rangle$, corresponding to energy the energies s and 0 , respectively. We assume that A initially is in state $|a_1\rangle$. By applying the operation $V_+(U_A)$ for $U_A := |a_0\rangle\langle a_1| + |a_1\rangle\langle a_0|$, we find that the state $\bar{\sigma}$ is translated up one rung along the ladder, to a new state σ_{out} . Hence, the system is again in a state with at least two quanta. Since these operations all have been performed on states safely away from the ground state, we can conclude that $\mathcal{C}(\sigma_{\text{out}}) = \mathcal{C}(\sigma_{\text{in}})$. By iterating this procedure we can thus conclude that the set of channels \mathcal{C} that can be reach by this reservoir, can be maintained for arbitrarily many transformations. (See Proposition 5 in Appendix B 3 for more details.)

Decay of coherence in the Jaynes-Cummings model.— Although for many theoretical purposes it is enough to have demonstrated the possibility to make coherence into a catalytic resource, it is nevertheless relevant to ask to what extent these phenomena are noticeable in more general systems, especially if one would consider experimental investigations of these questions. For this purpose, we consider the Jaynes-Cummings (JC) model [1, 2], and

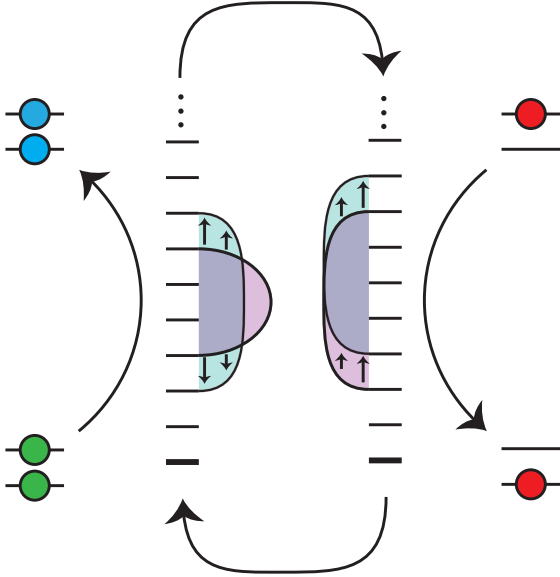


FIG. 2: **Regenerative catalytic cycles.** When the energy reservoir interacts with a two-level system (left part of the figure) the projection of its state onto the number states can increase with at most one level up and down in energy. As long as the projection onto the ground state of the reservoir remains zero, the set of channels that the reservoir can induce on the system stays intact from one interaction to the next. By using another two-level system in a pure excited state (right part of the figure) to inject energy into the reservoir, the state can be translated ‘rigidly’ up along the energy ladder by one step. By alternating every use of the reservoir with such a pumping, the state of the reservoir can be kept away from the ground state, thus maintaining the coherence properties indefinitely. Note that the state of the energy reservoir does change from one cycle to the next, e.g., the range of number states onto which it projects may become broader for each step. However, the *relevant aspects* of the state, which determines its capacity to induce channels, remains constant.

provide numerical evidence that, akin to the half-infinite ladder model, the life-time (in the ‘softer’ sense of decay rates) of the coherence of a fixed-width superposition can be made arbitrarily long by increasing its average energy (see Fig. 3 and Appendix C). The JC Hamiltonian, $H_{JC} = g\sigma_+ \otimes a + g\sigma_- \otimes a^\dagger$, is a common model for interactions between a single electromagnetic mode and a two-level system. Here a, a^\dagger are the standard annihilation and creation operators $[a, a^\dagger] = \hat{1}_E$, and $\sigma_+ = |\psi_1\rangle\langle\psi_0|$, $\sigma_- = |\psi_0\rangle\langle\psi_1|$. As described in figure 3, we measure, via the fidelity, the capacity of the reservoir to put a sequence of two-level systems into superpositions of their energy levels. As seen, the decay of the fidelity appears to become slower the further up in energy the initial state is. (For more details, see Appendix C 2.)

Application: Coherence in expected work extraction.—

In the work extraction setup we are given a system in a state characterized by a density operator ρ and a Hamiltonian H_S . We are furthermore given access to a heat

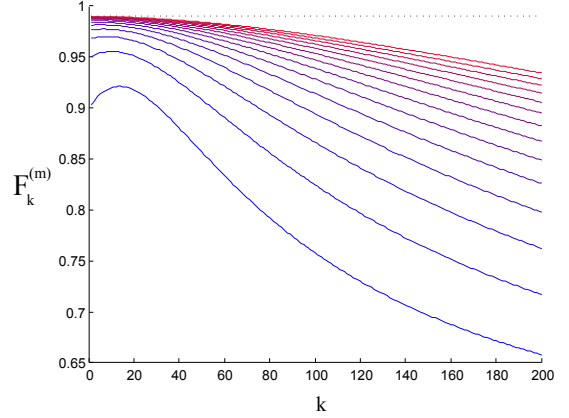


FIG. 3: **‘Lifetime’ of coherence in the Jaynes-Cummings model.** These graphs depicts the decay of the fidelity by which the superposition $|\phi\rangle = (|\psi_0\rangle - i|\psi_1\rangle)/\sqrt{2}$ can be created from the ground state $|\psi_0\rangle$, in a collection of two-level systems that sequentially interact with one single reservoir, according to the JC model for a fixed time-step. Each curve corresponds to a different initial state of the reservoir, and shows the square of the fidelity, $F_k^{(m)} = \langle\phi|\Phi(\sigma_m^{(k)})|\phi\rangle$, against the number of times k the reservoir has been used. Each initial state $\sigma_m^{(0)} = |\eta_m\rangle\langle\eta_m|$ of the reservoir is a uniform superposition $|\eta_m\rangle = \sum_{l=-24}^{25} |l_0(m) + l\rangle$, where $l_0(m) = (4m + 1)^2$, for $m = 4, 6, 8, \dots, 28, 30$, counted from below on the right hand side of the figure. These graphs suggest that the lifetime (counted in number of iterations) of the coherence can be made arbitrarily long merely by increasing the energy of the state, without changing the ‘width’ of the initial superposition. This behaviour echoes the properties of the half-infinite ladder model. The dotted line is the value of $(1 + |\langle\eta_m|\Delta|\eta_m\rangle|)/2 = 0.99$, which would be the fidelity reached in the doubly-infinite ladder-model for these initial states.

bath of temperature T , and our task is to extract as much useful energy as possible by equilibrating S with respect to the heat bath. Here we consider the question of how much work that can be extracted in an average sense, i.e., we strive to optimize the expected yield. Standard results [5, 34–36] suggest that the expected work content of a system is characterized by

$$\mathcal{A}_{\text{standard}}(\rho, H_S) = kTD(\rho\|G(H_S)), \quad (9)$$

where $D(\rho\|\eta) = \text{Tr}(\rho \ln \rho) - \text{Tr}(\rho \ln \eta)$ is the relative von-Neumann entropy, k is Boltzmann’s constant, $G(H_S) = \exp(-\beta H_S)/Z(H_S)$, $Z(H_S) = \text{Tr} \exp(-\beta H_S)$, and $\beta = 1/(kT)$. (With a base 2 logarithm, \log_2 , in the relative entropy, we would get an extra factor $\ln(2)$ in (9).) In models that do not include an explicit energy reservoir one can confirm (see Appendix D) that the optimal expected work content indeed can be characterized by equation (9).

However, in light of the above considerations on coherence, such approaches appears to implicitly assume the access to some form of ideal coherence resources. As men-

tioned earlier, it has been shown [19] that without access to any coherence, the optimal expected work content is not characterized by (9), but rather by

$$\mathcal{A}_{\text{diagonal}}(\rho, H_S) = kTD([\rho]_{H_S} \| G(H_S)), \quad (10)$$

where $[\rho]_{H_S} = \sum_l P_l \rho P_l$, where P_l are the projectors onto the eigenspaces of H_S . The operation $[\cdot]_{H_S}$ is an example of a ‘pinching’, where we here remove all off-diagonal blocks with respect to energy eigenspaces of the Hamiltonian. However, one can show (Appendix E) that access to coherence increase the amount of work that can be extracted. Furthermore, in the limit of very large degree of coherence (large L for the energy reservoir states $|\eta_{L,l_0}\rangle$), one regains equation (9). (See Appendix E3.) From these observations we can conclude that coherence sails up as an important resource alongside with expected work content in a quantum thermodynamic setting. The question is how they relate to each other; how much of one resource can be gained by spending the other? The fact that we here perform the work-extraction analysis entirely within the doubly-infinite ladder model, implies that we only use the coherence catalytically. In other words, we do not ‘spend’ any coherence at all.

Coherence and single-shot work extraction.— We have here focused on expected energy gain in work extraction. However, these results nevertheless suggest a couple of observations in relation to single-shot work extraction. As observed in the single-shot setting [12–18], the expected work content may not always correspond to ordered ‘work-like’ energy (see discussions in [14]). Although the coherence is preserved in the dynamics of the doubly-infinite ladder-model, the state nevertheless appears to gradually get more broadly distributed over the ladder (see Appendix A 6 for a simple example). Due to this one may suspect that the ‘quality’ of the energy decreases. This observation maybe gets even more pronounced for the half-infinite ladder model, where we inject energy in the form of pure excited states to maintain the coherence. Since such pure excited states correspond to very ordered energy, the spreading of the distribution again suggest a degradation of the quality of the energy. (One may note that in terms of expected energy, the injection is not lost, but merely adds to the extracted energy.) These observations raise the question whether there is a cost in terms of ordered energy (as opposed to expected energy) associated with catalytic coherence. If one would compare models that do and do not preserve coherence, would there in some sense be different costs of noise-free energy for performing operations? Such investigations would require the development of a (preferably operational) measure of ordered energy that takes into account coherence in the energy carrier *per se*.

Conclusions and outlook.— Here we have shown that there exists models where coherence is a catalytic resource. These models consists of a specific choice of Hamiltonian for a reservoir that carries energy and coherence. Furthermore, they specify rules on the allowed dynamics in the interactions between a target-system and

the reservoir. Here we have focused on the set of channels that the reservoir can induce on the system via the allowed interactions, and demonstrated that the more coherence there is in the reservoir, the more coherent the induced operations become. In the ‘doubly-infinite ladder model’ we have proved that the set of induced channels is invariant under repeated applications of the reservoir, and in this operational sense, the coherence resource in the reservoir is thus catalytic. To remedy the problem that the spectrum of this model is unbounded from below, we introduce the ‘half-infinite ladder’. For certain classes of initial states of the reservoir, and a procedure where energy is regularly injected into the reservoir, this model simulates the doubly-infinite ladder, and is thus also catalytic. We have furthermore made a brief numerical analysis of the Jaynes-Cummings model, which indicates a behaviour that is reminiscent of the more idealized models. As an application of these findings, we have analyzed work extraction on non-equilibrium states against a fixed-temperature heat bath, and shown that a catalytic use of coherence can release all the expected work content of a system by individual operations.

As mentioned above, we have here analyzed the coherence in the reservoir in terms of the channels that the reservoir can induce. For the models we have considered, the ‘relevant’ aspects of the state of the reservoir are the expectation values $\langle \Delta^a \rangle$. This certainly depends on the ‘allowed dynamics’ we have declared in the models, and thus different models could potentially single out different ‘types’ of coherence as ‘relevant’, and potentially invariant. How this should be understood in a wider perspective requires further analysis. As maybe already has been indicated, the question of coherence is closely related to the concept of reference frames and symmetry preserving operations [21–23]. The considerable machinery that has been developed in this context may favorably be applied to further analyze the question of catalytic coherence, but also to provide alternative ways to analyze these questions.

We have here argued that coherence is an important thermodynamic quantity. One may thus ask why we normally do not encounter coherence in discussions on thermodynamics. In the limit of many independent subsystems in identical states, and identical non-interacting Hamiltonians, the global state has a low degree of ‘off-diagonality’ with respect to the joint energy eigenspaces [10, 15, 19] and hence the difference between (9) and (10) becomes small. Since many realistic systems can be argued to be ‘close’ to this situation, we would thus normally not expect to see the effects of coherence. However, as demonstrated above, this is no longer true when we step outside of this ‘multi-copy’ setting. To obtain a complete theory of quantum thermodynamics, this is arguably a necessary step.

Acknowledgments

The author thanks Lidia del Rio, Philippe Faist, and Paul Skrzypczyk for useful discussions. This research

was supported by the Excellence Initiative of the German Federal and State Governments (grant ZUK 43).

Appendix A: Doubly infinite energy ladder

As a first step to investigate the catalytic properties of coherence, we here consider a simple model for energy conserving operations on a combination of a system and an energy reservoir. This model has previously been considered in [19].

1. The doubly-infinite ladder model

Our model consists of two parts: A ‘system’ S , modeled by a finite-dimensional complex Hilbert space \mathcal{H}_S , and an ‘energy reservoir’ E , modeled by a separable Hilbert space \mathcal{H}_E . We let $\{|j\rangle\}_{j \in \mathbb{Z}}$ denote a complete orthonormal basis of \mathcal{H}_E . Here \mathbb{Z} denote the set of (negative as well as positive) integers. Given a real number $s > 0$ we define the following unbounded Hermitian operator

$$H_E^s := s \sum_{j \in \mathbb{Z}} j |j\rangle \langle j|. \quad (\text{A1})$$

We will in the following regard this operator as the Hamiltonian of the energy reservoir E , although, needless to say, it is somewhat unphysical, since it has a ‘bottomless’ spectrum. However, we will remedy this problem in Section B. ([19] mostly uses a continuum reservoir rather than a discrete model. Although see the remarks in Appendix G of [19].)

Given a finite-dimensional Hilbert space \mathcal{H}_S and a real number $s > 0$ we let $H_s(\mathcal{H}_S)$ denote the set of Hermitian operators on \mathcal{H}_S such that all its eigenvalues are integer multiples of s , i.e., all its eigenvalues can be written as $h_n = sz_n$ where $z_n \in \mathbb{Z}$. The reason for this construction is to ‘match’ all energy differences in H_S with the energy spacings in $H_E^{(s)}$. By this we make sure that for each pair of eigenstates in S there is a transition (in fact infinitely many) in E that can compensate for the change in energy. The eigenspaces of the the joint system $H_S + H_E^{(s)}$ (we should strictly speaking write $H_S \otimes \hat{1}_E + \hat{1}_S \otimes H_E^{(s)}$) have a very simple structure.

Lemma 1. *Let $s > 0$ be given. Let \mathcal{H}_S be a finite-dimensional Hilbert space with $\dim \mathcal{H}_S = N$, $H_S \in H_s(\mathcal{H}_S)$, with eigenvectors $|\psi_n\rangle$ and corresponding eigen-*

values sz_n (where $z_n \in \mathbb{Z}$). Then

$$H_S + H_E^{(s)} = s \sum_{j \in \mathbb{Z}} j P^{(j)}, \quad (\text{A2})$$

$$P^{(j)} := \sum_{n=1}^N |\psi_n\rangle \langle \psi_n| \otimes |j - z_n\rangle \langle j - z_n|$$

is an eigenvalue decomposition, where $P^{(j)}$ is the projector onto the eigenspace corresponding to the eigenvalue js .

The whole purpose of the construction with the energy reservoir is to keep track of energy. Hence, for the dynamics on the joint system SE we consider unitary operations that conserve the total energy, i.e., that commute with the joint Hamiltonian $H_S + H_E^{(s)}$. However, we furthermore demand that the dynamics should act ‘uniformly’ over the energy levels in the reservoir. To express this a bit more precisely, we introduce a unitary operator Δ on \mathcal{H}_E , which can be thought of as a ‘rigid translation’ along the doubly-infinite energy ladder.

$$\Delta := \sum_{k \in \mathbb{Z}} |k+1\rangle \langle k|. \quad (\text{A3})$$

One can see that

$$\Delta^a \Delta^b = \Delta^b \Delta^a = \Delta^{a+b}, \quad \Delta^{a\dagger} = \Delta^{-a}, \quad a, b \in \mathbb{Z}, \quad (\text{A4})$$

where we let $\Delta^0 = \hat{1}_E$.

Definition 1. *Let $s > 0$, and let \mathcal{H}_S be a finite-dimensional Hilbert space and let $H_S \in H_s(\mathcal{H}_S)$. We use the following notation*

- $L(\mathcal{H}_S)$ is the set of linear operators on \mathcal{H}_S , and $U(\mathcal{H}_S)$ the set of unitary operators on \mathcal{H}_S .
- $\mathbb{LT}(H_S + H_E^{(s)})$ is the set of bounded linear operators X on $\mathcal{H}_S \otimes \mathcal{H}_E$ such that $[X, H_S + H_E^{(s)}] = 0$, and $[X, \hat{1}_S \otimes \Delta^a] = 0$ for all $a \in \mathbb{Z}$.
- $\mathbb{UT}(H_S + H_E^{(s)})$ is the set of unitary operators U on $\mathcal{H}_S \otimes \mathcal{H}_E$ such that $[U, H_S + H_E^{(s)}] = 0$, and $[U, \hat{1}_S \otimes \Delta^a] = 0$ for all $a \in \mathbb{Z}$.
- Given a Hilbert space \mathcal{H} we let $\mathcal{S}(\mathcal{H})$ denote the set of positive semi-definite trace class operators with trace 1 on \mathcal{H} , which we refer to as density operators or ‘states’. We let $\mathcal{S}_+(\mathcal{H})$ denote the set of density operators with full support.

The set $\mathbb{UT}(H_S + H_E^{(s)})$ is precisely the set of evolution operators we allow in our model; they conserve energy, and commute with all energy translations Δ^a . The following lemma shows that we can describe the set $\mathbb{UT}(H_S + H_E^{(s)})$ in a very convenient manner.

Lemma 2. *Let \mathcal{H}_S be finite-dimensional with $N := \dim \mathcal{H}_S$, and let $H_S \in H_s(\mathcal{H}_S)$. Let $|\psi_n\rangle$ be eigenvectors of H_S with corresponding eigenvalues sz_n . Define the mapping*

$$V(Q) := \sum_{n,n'=1}^N |\psi_n\rangle\langle\psi_n|Q|\psi_{n'}\rangle\langle\psi_{n'}| \otimes \Delta^{z_{n'}-z_n}, \quad (\text{A5})$$

for all Q in $L(\mathcal{H}_S)$. The mapping V is a bijection between $L(\mathcal{H}_S)$ and $\mathbb{LT}(H_S + H_E^{(s)})$, which preserves the structures of these operator spaces, in the sense that

$$\begin{aligned} V(\alpha A + \beta B) &= \alpha V(A) + \beta V(B), \\ V(AB) &= V(A)V(B), \\ V(A^\dagger) &= V(A)^\dagger, \\ V(\hat{1}_S) &= \hat{1}_S \otimes \hat{1}_E, \end{aligned} \quad (\text{A6})$$

for all $A, B \in L(\mathcal{H}_S)$ and all $\alpha, \beta \in \mathbb{C}$. As a consequence, V is a bijection between $U(\mathcal{H}_S)$ and $\mathbb{UT}(H_S + H_E^{(s)})$.

Equation (A5) provides us with a very convenient handle on $\mathbb{UT}(H_S + H_E^{(s)})$, since we can reach each element of $\mathbb{UT}(H_S + H_E^{(s)})$ uniquely by $V(U)$ for some $U \in U(\mathcal{H}_S)$. Due to this reason we will in this investigation hardly ever refer explicitly to the set $\mathbb{UT}(H_S + H_E^{(s)})$, but rather describe it indirectly via V .

It is worth noting that one can rewrite the mapping V as

$$\begin{aligned} V(Q) &= \sum_{j \in \mathbb{Z}} V_j(Q), \\ V_j(Q) &:= \sum_{n,n'=1}^N |\psi_n\rangle\langle\psi_n|Q|\psi_{n'}\rangle\langle\psi_{n'}| \otimes |j - z_n\rangle\langle j - z_{n'}|, \end{aligned} \quad (\text{A7})$$

where $V_j(Q) = P^{(j)}V_j(Q) = V_j(Q)P^{(j)}$. Hence, what V essentially does is to embed an infinite number of ‘copies’ of Q into the space of operators on $\mathcal{H}_S \otimes \mathcal{H}_E$. Moreover, if $U \in U(\mathcal{H}_S)$ then $V_j(U)V_j(U)^\dagger = V_j(U)^\dagger V_j(U) = P^{(j)}$.

Proof of Lemma 2. Since $V(Q)$ merely is a finite sum of bounded operators, it follows that V maps $L(\mathcal{H}_S)$ into the set of bounded linear operators on $\mathcal{H}_S \otimes \mathcal{H}_E$. Since all Δ^a commutes, it follows that $[V(Q), \hat{1}_S \otimes \Delta^a] = 0$ for all $a \in \mathbb{Z}$. One can also verify (e.g. via showing $[H_E^{(s)}, \Delta^a] = sa\Delta^a$) that $[V(Q), H_S + H_E^{(s)}] = 0$. Hence $V(L(\mathcal{H}_S)) \subseteq \mathbb{LT}(H_S + H_E^{(s)})$. It furthermore only requires a verifications to see that the conditions in (A6) hold. It thus only remains to show that V is surjective

[i.e., $V(L(\mathcal{H}_S)) = \mathbb{LT}(H_S + H_E^{(s)})$] and injective (i.e., two different elements in $L(\mathcal{H}_S)$ cannot get mapped to the same element). The injectivity can be obtained if we apply (A5) to $V(Q_1) = V(Q_2)$ and note that Δ^a is never the zero operator. For the surjectivity we assume $Y \in \mathbb{LT}(H_S + H_E^{(s)})$, and note that $[Y, H_S + H_E^{(s)}] = 0$ implies that $Y = \sum_{j \in \mathbb{Z}} P^{(j)}Y P^{(j)}$, with $P^{(j)}$ as in Lemma 1. Hence,

$$Y = \sum_j \sum_{n,n'} (|\psi_n\rangle\langle j - z_n|)Y(|\psi_{n'}\rangle\langle j - z_{n'}|) |\psi_n\rangle\langle\psi_{n'}| \otimes |j - z_n\rangle\langle j - z_{n'}|. \quad (\text{A8})$$

Using $|j - z_{n'}\rangle = \Delta^j| - z_{n'}\rangle$, and $[\hat{1}_S \otimes \Delta^{\dagger j}]Y[\hat{1}_S \otimes \Delta^j] = Y$ (since $[Y, \hat{1}_S \otimes \Delta^j] = 0$) we can conclude that

$$X := \sum_{n,n'} (|\psi_n\rangle\langle -z_n|)Y(|\psi_{n'}\rangle\langle -z_{n'}|) |\psi_n\rangle\langle\psi_{n'}|, \quad (\text{A9})$$

is such that $Y = V(X)$, which proves the Lemma. Combining the bijectivity of V with (A6) yields that V also is a bijection between $U(\mathcal{H}_S)$ and $\mathbb{UT}(H_S + H_E^{(s)})$. \square

2. Induced channels

Given the unitary operators $V(U)$ on the joint system SE we can induce channels on S and E .

For a fixed state σ on the energy reservoir, each choice of unitary U on S induces a channel $\Phi_{\sigma,U}^S$ on system S via

$$\begin{aligned} \Phi_{\sigma,U}^{S,H}(\rho) &:= \text{Tr}_E V(U)\rho \otimes \sigma V(U)^\dagger \\ &= \sum_{n,n',m,m'=1}^N \langle\psi_n|U|\psi_{n'}\rangle\langle\psi_{n'}|\rho|\psi_{m'}\rangle\langle\psi_{m'}|U^\dagger|\psi_n\rangle \\ &\quad |\psi_n\rangle\langle\psi_m| \text{Tr}(\Delta^{z_{n'}-z_n-z_{m'}+z_m}\sigma). \end{aligned} \quad (\text{A10})$$

We denote the set of channels that can be reached on S , by using σ as a resource, by

$$\mathcal{C}^{S,H}(\sigma) := \left\{ \Phi_{\sigma,U}^{S,H} : U \in U(\mathcal{H}_S) \right\}. \quad (\text{A11})$$

As seen, both $\Phi_{\sigma,U}^{S,H}$ and $\mathcal{C}^{S,H}$ depend on the choice of Hamiltonian H_S , both via the eigenstates $|\psi_n\rangle$, and the eigenvalues sz_n . However, we will in the following often drop one or both of the superscripts when they are not needed, i.e., $\Phi_{\sigma,U}^{S,H}$, $\Phi_{\sigma,U}^S$, and $\Phi_{\sigma,U}$ denote the same object. Similarly for $\mathcal{C}^{S,H}$, \mathcal{C}^S , and \mathcal{C} .

We can also express the effect of $V(U)$ on the energy reservoir via the channel

$$\begin{aligned} \Lambda_{\rho,U}(\sigma) &:= \text{Tr}_S V(U)\rho \otimes \sigma V(U)^\dagger \\ &= \sum_{n,n',m'=1}^N \langle\psi_n|U|\psi_{n'}\rangle\langle\psi_{n'}|\rho|\psi_{m'}\rangle\langle\psi_{m'}|U^\dagger|\psi_n\rangle \\ &\quad \Delta^{z_{n'}-z_n}\sigma\Delta^{z_{m'}-z_n}{}^\dagger. \end{aligned} \quad (\text{A12})$$

3. Approximation of energy-mixing unitary operations

At first sight it may not be clear what kind of operations one can reach on S . Here we show that in the limit of large coherence, all unitary operators can be implemented on S . As mentioned in the main text, this model can be viewed as a simplified version of transitions in atoms induced by coherent electromagnetic fields (see e.g. Chapter 7 in [33]). In both cases the key-point is that the state of the energy reservoir is such that adding or removing a small amount of energy does not change it much.

Define the states

$$|\eta_{L,l_0}\rangle := \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} |l + l_0\rangle, \quad (\text{A13})$$

i.e., this is a uniform superposition of $L \geq 1$ sequential energy eigenstates, starting at l_0 . One can verify that

$$\langle \eta_{L,l_0} | \Delta^a | \eta_{L,l_0} \rangle = \max\left(0, 1 - \frac{|a|}{L}\right). \quad (\text{A14})$$

In other words, the shifted $\Delta^a |\eta_{L,l_0}\rangle$ and the unshifted state $|\eta_{L,l_0}\rangle$ are almost indistinguishable if $L \gg a$.

In the following, we shall see that the difference between the induced channel $\Phi_{\sigma,U}$ and the unitary channel

$$\mathcal{U}_U(\rho) := U\rho U^\dagger, \quad (\text{A15})$$

becomes very small when $L \gg a$. Hence, in the limit of large L , we can implement arbitrary unitary operations on S .

To this end we do, on the set of super-operators $\Gamma : L(\mathcal{H}_S) \rightarrow L(\mathcal{H}_S)$, consider the operator norm induced by the trace norm $\|\Gamma\|_{\text{tr}} := \sup_{\|X\|_1=1} \|\Gamma(X)\|_1$. By trivially extending the super-operator to a sufficiently large ancillary Hilbert space \mathcal{H}_A (it is sufficient with $\dim \mathcal{H}_A = \dim \mathcal{H}_S$ [37]) one can define the diamond norm [37, 38] $\|\Gamma\|_\diamond := \|\Gamma \otimes I_A\|_{\text{tr}}$, where I_A is the identity map on $L(\mathcal{H}_A)$.

In the following, we let $\mathbb{C}^{n \times n}$ denote the set of $n \times n$ matrices with complex entries. Furthermore, for $A, B \in \mathbb{C}^{n \times n}$ we let $A * B$ denote the Hadamard product, i.e. the component-wise product $(A * B)_{jk} := A_{jk} B_{jk}$. We furthermore let $\|\cdot\|$ denote the spectral norm $\|Q\| := \sup_{\|\psi\|=1} \|Q\psi\|$. For an arbitrary $Q \in \mathbb{C}^{N \times N}$ we let in the following $(s_l(Q))_{l=1}^N$ denote the singular values of Q (including zero values) ordered non-increasingly $s_1(Q) \geq s_2(Q) \geq \dots \geq s_N(Q)$.

Lemma 3 (Theorem 5.5.4 on p. 334 in [39]). *Let $A, B \in \mathbb{C}^{N \times N}$ then*

$$\sum_{l=1}^k s_l(A * B) \leq \sum_{l=1}^k s_l(A) s_l(B), \quad (\text{A16})$$

for $k = 1, \dots, N$.

Since $\|Q\|_1 = \sum_{l=1}^N s_l(Q)$ and $\|Q\| = s_1(Q)$ it follows directly from the above lemma that

$$\|A * B\|_1 \leq \|A\| \|B\|_1. \quad (\text{A17})$$

Lemma 4 (Eq. (3.7.2) on p. 223 in [39]). *Let $C \in \mathbb{C}^{n \times n}$, then*

$$\|C\| \leq \sqrt{\max_j \sum_k |C_{jk}|} \sqrt{\max_k \sum_j |C_{jk}|}. \quad (\text{A18})$$

Proposition 1. *Let $s > 0$, \mathcal{H}_S with $\dim \mathcal{H}_S = N$, and $H_S \in \mathcal{H}_s(\mathcal{H}_S)$, with eigenvalues sz_n and corresponding orthonormal eigenvectors $|\psi_n\rangle$. Let $\Phi_{\sigma,U}$ be as defined in equation (A10), and let $|\eta_{L,l_0}\rangle$ be as in equation (A13), for $l_0 \in \mathbb{Z}$ and $L \in \mathbb{N}$. Let $z_{\max} := \max_n z_n$ and $z_{\min} := \min_n z_n$. If*

$$L \geq 2(z_{\max} - z_{\min}), \quad (\text{A19})$$

then for every $U \in U(\mathcal{H}_S)$,

$$\|\mathcal{U}_U - \Phi_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|,U}\|_{\text{tr}} \leq 2N^2 \frac{z_{\max} - z_{\min}}{L} \quad (\text{A20})$$

and

$$\|\mathcal{U}_U - \Phi_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|,U}\|_\diamond \leq 2N^3 \frac{z_{\max} - z_{\min}}{L}. \quad (\text{A21})$$

Proof. Let \mathcal{H}_A be a Hilbert space with orthonormal basis $\{|a_n\rangle\}_{n=1}^M$. (The choice of M will in the end give us the induced norm ($M = 1$) or the diamond norm ($M = N$).) Let $X \in L(\mathcal{H}_S \otimes \mathcal{H}_A)$ be such that $\|X\|_1 = 1$. For the sake of compactness we will use the notation $|\psi_k, a_n\rangle := |\psi_k\rangle|a_n\rangle$.

Due to $L \geq 2(z_{\max} - z_{\min})$ it follows that $|z_k - z_{k'} - z_l + z_{l'}| \leq L$. Combined with equations (A10) and (A14) this yields

$$\begin{aligned} & [U \otimes \hat{1}_A] X [U^\dagger \otimes \hat{1}_A] - [\Phi_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|,U}^S \otimes I_A](X) \\ &= \sum_{k,k',l,l'} \sum_{n,n'} \langle \psi_{k'} | U | \psi_k \rangle \langle \psi_k, a_n | X | \psi_l, a_{n'} \rangle \langle \psi_l | U^\dagger | \psi_{l'} \rangle \\ & \quad |\psi_{k'}, a_n\rangle \langle \psi_{l'}, a_{n'}| \frac{|z_k - z_{k'} - z_l + z_{l'}|}{L} \end{aligned} \quad (\text{A22})$$

By the triangle inequality for the trace norm

$$\begin{aligned} & \left\| [U \otimes \hat{1}_A] X [U^\dagger \otimes \hat{1}_A] - [\Phi_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|,U}^S \otimes I_A](X) \right\|_1 \\ & \leq \sum_{n,n'} \sum_{k,l} \|A^{(k,l,n,n')} * B^{(k,l)}\|_1, \end{aligned} \quad (\text{A23})$$

where ‘ $*$ ’ denotes the Hadamard product, and where the matrices $A^{(k,l,n,n')} := [A_{k',l'}^{(k,l,n,n')}]_{k',l'}$ and $B^{(k,l)} := [B_{k',l'}^{(k,l)}]_{k',l'}$ are defined as

$$\begin{aligned} A_{k',l'}^{(k,l,n,n')} &:= \langle \psi_{k'} | U | \psi_k \rangle \langle \psi_k, a_n | X | \psi_l, a_{n'} \rangle \langle \psi_l | U^\dagger | \psi_{l'} \rangle \\ B_{k',l'}^{(k,l)} &:= \frac{|z_k - z_{k'} - z_l + z_{l'}|}{L}. \end{aligned} \quad (\text{A24})$$

By Lemma 4, $\|B^{(k,l)}\| \leq N(z_{\max} - z_{\min})/L$. One can confirm $\|A^{(k,l,n,n')}\|_1 = |\langle \psi_k, a_n | X | \psi_l, a_{n'} \rangle|$. By combining these observations with equations (A17) and (A23) we can conclude that

$$\begin{aligned} & \left\| [U \otimes \hat{I}_A] X [U^\dagger \otimes \hat{I}_A] - [\Phi_{\eta_{L,l_0}^S}^S]_{\langle \eta_{L,l_0}^S |, U \otimes I_A} (X) \right\|_1 \\ & \leq N \frac{2(z_{\max} - z_{\min})}{L} \sum_{n,n'} \sum_{k,l} |\langle \psi_k, a_n | X | \psi_l, a_{n'} \rangle| \\ & \leq MN^2 \frac{2(z_{\max} - z_{\min})}{L}, \end{aligned} \quad (\text{A25})$$

where the last inequality follows by Cauchy-Schwarz. If we choose the ancillary Hilbert space to be trivial, i.e., $\dim \mathcal{H}_A = M = 1$, then we obtain the bound on the norm $\|\cdot\|_{\text{tr}}$. It is sufficient to choose the ancillary Hilbert space to have the same dimension as \mathcal{H}_S to obtain the diamond norm [37]. In this case $M = N$ and we obtain the corresponding bound. \square

4. A catalytic property

The previous section shows that the set of channels $\mathcal{C}(\sigma)$ which we can reach on system S depends on the coherence in the state σ of the energy reservoir E . In other words, the coherence in the state of E is a resource for the implementation of channels on S . Here we show that this resource is catalytic, i.e., if we use E to implement a channel on one system S_1 , then this does not diminish the capacity of E to subsequently implement channels on another system S_2 . It is worth noting that this catalytic property holds for all states σ ; it is not limited to states with a large degree of coherence.

Consider the expression for the channel $\Phi_{\sigma,U}^S$ in equation (A10). As one can see, the only way in which $\Phi_{\sigma,U}^S$ depends on σ , is via the expectation value of powers of Δ , i.e., on the sequence of numbers $(\text{Tr}(\Delta^a \sigma))_{a \in \mathbb{Z}}$. (Strictly speaking, it only depends on a finite number of them.) One should note that this set of numbers does not uniquely define σ . If one think of σ represented as a matrix in the energy eigenbasis, i.e., in terms of the matrix $[\langle j | \sigma | j' \rangle]_{j,j'}$, one can see that $\text{Tr}(\Delta^0 \sigma)$ corresponds to the trace, $\text{Tr}(\Delta^1 \sigma) = \sum_j \langle j | \sigma | j+1 \rangle$ to the sum of the elements in the first upper diagonal, $\text{Tr}(\Delta^2 \sigma)$ to the second, etc. The following lemma shows that these expectation values are invariant under the action of the channel $\Lambda_{\rho,\sigma}$ on the energy reservoir.

Lemma 5. *Let $s > 0$, \mathcal{H}_S finite-dimensional, and $H_S \in H_s(\mathcal{H}_S)$. Then*

$$\text{Tr}(\Delta^a \Lambda_{\rho,U}(\sigma)) = \text{Tr}(\Delta^a \sigma), \quad (\text{A26})$$

for all $\sigma \in \mathcal{S}(\mathcal{H}_E)$, $U \in U(\mathcal{H}_S)$, $\rho \in \mathcal{S}(\mathcal{H}_S)$, and $a \in \mathbb{Z}$, where $\Lambda_{\rho,U}$ is as defined in equation (A12).

Proof. Let sz_k and $|\psi_k\rangle$ be the eigenvalues and corresponding eigenvectors of H_S , and put $U_{k',k} := \langle \psi_{k'} | U | \psi_k \rangle$. Then

$$\begin{aligned} & \text{Tr}(\Delta^a \Lambda_{\rho,U}(\sigma)) \\ & = \sum_{k,l,k'} U_{k',k} U_{k',l}^* \langle \psi_k | \rho | \psi_l \rangle \text{Tr}(\Delta^a \Delta^{z_k - z_{k'}} \sigma \Delta^{z_l - z_{k'}}^\dagger) \\ & = \sum_{k,l,k'} U_{k',k} U_{k',l}^* \langle \psi_k | \rho | \psi_l \rangle \text{Tr}(\Delta^{-z_l + z_{k'}} \Delta^a \Delta^{z_k - z_{k'}} \sigma) \\ & = \sum_{k,l,k'} U_{k',k} U_{k',l}^* \langle \psi_k | \rho | \psi_l \rangle \text{Tr}(\Delta^{a+z_k - z_l} \sigma) \\ & = \sum_{k,l} \delta_{k,l} \langle \psi_k | \rho | \psi_l \rangle \text{Tr}(\Delta^{a+z_k - z_l} \sigma) \\ & = \sum_k \langle \psi_k | \rho | \psi_k \rangle \text{Tr}(\Delta^a \sigma) \\ & = \text{Tr}(\Delta^a \sigma). \end{aligned} \quad (\text{A27})$$

\square

The following proposition tells us that if we use the reservoir twice, then the set of operations we can reach in the second application is independent of what we did in the first application. In other words, it is as if the first application had not happened. It is maybe also worth to point out that the only assumptions we impose on the Hamiltonians of the subsequent subsystems are that they are non-interacting, and satisfy the ‘energy matching condition’. We do not assume that they are identical (or even operating on Hilbert spaces of the same dimension). It is straightforward to generalize this proposition to more than two subsystems.

Proposition 2 (Catalytic coherence). *Let $s > 0$, and let \mathcal{H}_{S_1} and \mathcal{H}_{S_2} be finite-dimensional Hilbert spaces. Let $H_{S_1} \in H_s(\mathcal{H}_{S_1})$ and $H_{S_2} \in H_s(\mathcal{H}_{S_2})$. Let $U_1 \in U(\mathcal{H}_{S_1})$ and $\rho_1 \in \mathcal{S}(\mathcal{H}_{S_1})$, $U_2 \in U(\mathcal{H}_{S_2})$ and $\rho_2 \in \mathcal{S}(\mathcal{H}_{S_2})$. Let Λ_{ρ_1,U_1} be the channel on E as defined by equation (A12) for interaction with system S_1 , and let Φ^{S_2} be the channel as defined by equation (A10), for system S_2 . Then*

$$\Phi_{\Lambda_{\rho_1,U_1}(\sigma), U_2}^{S_2, H_{S_2}}(\rho_2) = \Phi_{\sigma, U_2}^{S_2, H_{S_2}}(\rho_2). \quad (\text{A28})$$

Equivalently, this can be reformulated as

$$\mathcal{C}^{S_2, H_{S_2}}(\Lambda_{\rho_1,U_1}(\sigma)) = \mathcal{C}^{S_2, H_{S_2}}(\sigma). \quad (\text{A29})$$

Proof. Let $|\psi_n^{(1)}\rangle$ and $sz_n^{(1)}$ be orthonormal eigenvectors and corresponding eigenvalues of H_{S_1} , and let $|\psi_m^{(2)}\rangle$ and $sz_m^{(2)}$ be orthonormal eigenvectors and corresponding eigenvalues of H_{S_2} . By equation (A10) we can write

$$\begin{aligned} & \Phi_{\Lambda_{\rho_1,U_1}(\sigma), U_2}^{S_2, H_{S_2}}(\rho_2) \\ & = \sum_{k,k',l,l'} \langle \psi_{k'}^{(2)} | U_2 | \psi_k^{(2)} \rangle \langle \psi_k^{(2)} | \rho_2 | \psi_l^{(2)} \rangle \langle \psi_l^{(2)} | U_2^\dagger | \psi_{l'}^{(2)} \rangle \\ & \quad |\psi_{k'}^{(2)}\rangle \langle \psi_{l'}^{(2)} | \text{Tr}[\Delta^{z_k^{(2)} - z_{k'}^{(2)} - z_l^{(2)} + z_{l'}^{(2)}} \Lambda_{\rho_1,U_1}(\sigma)]. \end{aligned} \quad (\text{A30})$$

By Lemma 5 we know that

$$\begin{aligned} \text{Tr}[\Delta^{z_k^{(2)}-z_{k'}^{(2)}-z_l^{(2)}+z_{l'}^{(2)}} \Lambda_{\rho_1, U_1}(\sigma)] \\ = \text{Tr}[\Delta^{z_k^{(2)}-z_{k'}^{(2)}-z_l^{(2)}+z_{l'}^{(2)}} \sigma], \end{aligned} \quad (\text{A31})$$

which proves (A28) and (A29). \square

In the above proposition we made no assumptions on the relation between the first and the second application of the reservoir. For the sake of clarity, we here also treat the special case that we do use the same Hamiltonian on both the first and second system.

Corollary 1. *Given the assumptions of Proposition 2, but with the additional assumptions $\mathcal{H}_{S_2} \approx \mathcal{H}_{S_1}$, $H := H_{S_2} = H_{S_1}$, it follows that*

$$\mathcal{C}^{S_2, H}(\Lambda_{\rho_1, U_1}(\sigma)) = \mathcal{C}^{S_1, H}(\sigma). \quad (\text{A32})$$

Strictly speaking the correct formulation of (A32) would be to say that the two sets of channels, $\mathcal{C}^{S_2, H}(\Lambda_{\rho_1, U_1}(\sigma))$ and $\mathcal{C}^{S_1, H}(\sigma)$, are isomorphic (since the channels act on different spaces).

Proof. Assume that \mathcal{H}_{S_2} and \mathcal{H}_{S_1} are isomorphic, with the isomorphism $X := \sum_{n=1}^N |\psi_n^{(2)}\rangle\langle\psi_n^{(1)}|$, where $H_{S_1}|\psi_n^{(1)}\rangle = z_n^{(1)}|\psi_n^{(1)}\rangle$, and $H_{S_2}|\psi_n^{(2)}\rangle = z_n^{(2)}|\psi_n^{(2)}\rangle$, with $z_n := z_n^{(2)} = z_n^{(1)}$ (and thus ' $H_{S_2} = H_{S_1}$ '). By equations (A30) and (A31) we see that $X^\dagger \Phi_{\Lambda_{\rho_1, U_1}(\sigma), U_2}^{S_2, H_{S_2}}(\rho_2)X = \Phi_{\sigma, X^\dagger U_2 X}^{S_1, H_{S_1}}(X^\dagger \rho_2 X)$. Hence, we can implement 'the same' channels on S_2 as we can on S_1 , and thus, in this sense, equation (A32) holds. \square

5. A stronger catalytic property

In the previous section we assumed that the system and the energy reservoir initially are uncorrelated (which indeed is necessary if we wish to express the evolution in terms of channels). Here we show that even if we initially have an arbitrary joint state on the reservoir and the systems, a catalytic property still holds.

Lemma 6. *Let $s > 0$, and let \mathcal{H}_{S_1} and \mathcal{H}_{S_2} be finite-dimensional Hilbert spaces. Let $H_{S_1} \in H_s(\mathcal{H}_{S_1})$ and $H_{S_2} \in H_s(\mathcal{H}_{S_2})$, and define $\mathcal{H}_S := \mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$ and $H_S := \hat{1}_2 \otimes H_{S_1} + H_{S_2} \otimes \hat{1}_{S_1}$. The mapping V , as defined in Lemma 2, can be defined on the spaces $\mathcal{H}_{S_1} \otimes \mathcal{H}_E^s$, $\mathcal{H}_{S_2} \otimes \mathcal{H}_E^s$, and $\mathcal{H}_S \otimes \mathcal{H}_E^s$, with respect to the Hamiltonians H_{S_1} , H_{S_2} , and H_S , respectively. For the sake of clarity we denote these with superscripts. It is the case that*

$$\begin{aligned} V^{H_S}(\hat{1}_{S_2} \otimes Q_1) &= \hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1), \quad \forall Q_1 \in L(\mathcal{H}_{S_1}), \\ V^{H_S}(Q_2 \otimes \hat{1}_{S_1}) &= V^{H_{S_2}}(Q_2) \otimes \hat{1}_{S_1}, \quad \forall Q_2 \in L(\mathcal{H}_{S_2}). \end{aligned} \quad (\text{A33})$$

Furthermore

$$\begin{aligned} [\hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1)][V^{H_{S_2}}(Q_2) \otimes \hat{1}_{S_1}] \\ = [V^{H_{S_2}}(Q_2) \otimes \hat{1}_{S_1}][\hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1)], \end{aligned} \quad (\text{A34})$$

for all $Q_1 \in L(\mathcal{H}_{S_1})$ and $Q_2 \in L(\mathcal{H}_{S_2})$.

Proof. Let $|\psi_n^{(1)}\rangle$ and $sz_n^{(1)}$ be orthonormal eigenvectors and corresponding eigenvalues of H_{S_1} , and let $|\psi_m^{(2)}\rangle$ and $sz_m^{(2)}$ be orthonormal eigenvectors and corresponding eigenvalues of H_{S_2} . For a compact notation we define $|\psi_m^{(2)}, \psi_n^{(1)}\rangle := |\psi_m^{(2)}\rangle|\psi_n^{(1)}\rangle$. By using equation (A5) we can write

$$\begin{aligned} V^{H_S}(\hat{1}_{S_2} \otimes Q_1) \\ = \sum_{nm, n'm'} \langle \psi_m^{(2)}, \psi_n^{(1)} | (\hat{1}_{S_2} \otimes Q_1) | \psi_{m'}^{(2)}, \psi_{n'}^{(1)} \rangle \\ \times |\psi_m^{(2)}, \psi_n^{(1)}\rangle \langle \psi_{m'}^{(2)}, \psi_{n'}^{(1)} | \otimes \Delta^{z_{n'}^{(1)}-z_n^{(1)}+z_{m'}^{(2)}-z_m^{(2)}} \\ = \hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1). \end{aligned}$$

An analogous reasoning holds for $V^{H_S}(Q_2 \otimes \hat{1}_{S_1})$. By Lemma 2 we know that V preserves operator multiplication, and thus

$$\begin{aligned} [\hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1)][V^{H_{S_2}}(Q_2) \otimes \hat{1}_{S_1}] \\ = V^{H_S}(\hat{1}_{S_2} \otimes Q_1) V^{H_S}(Q_2 \otimes \hat{1}_{S_1}) \\ = V^{H_S}(Q_2 \otimes Q_1) \\ = V^{H_S}(Q_2 \otimes \hat{1}_{S_1}) V^{H_S}(\hat{1}_{S_2} \otimes Q_1) \\ = [V^{H_{S_2}}(Q_2) \otimes \hat{1}_{S_1}][\hat{1}_{S_2} \otimes V^{H_{S_1}}(Q_1)]. \end{aligned}$$

\square

Suppose that we, apart from the energy reservoir E have two systems of interest S_1 and S_2 , and that the joint system $S_2 S_1 E$ initially is in some arbitrary joint state η .

One can compare two different scenarios. In the first scenario, we first implement a unitary operation $V^{H_{S_1}}$ on ES_1 followed by a unitary operation $V^{H_{S_2}}$ on ES_2 . In the second scenario, we only perform the second operation $V^{H_{S_2}}$ on ES_2 . The following proposition tells us that the reduced density operator on S_2 is the same in both scenarios.

Proposition 3 (Strong catalytic property). *Let $s > 0$, and let \mathcal{H}_{S_1} and \mathcal{H}_{S_2} be finite-dimensional Hilbert spaces. Let $H_{S_1} \in H_s(\mathcal{H}_{S_1})$ and $H_{S_2} \in H_s(\mathcal{H}_{S_2})$, and define $\mathcal{H}_S = \mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$ and $H_S = \hat{1}_2 \otimes H_{S_1} + H_{S_2} \otimes \hat{1}_{S_1}$. Let $U_1 \in U(\mathcal{H}_{S_1})$ and $U_2 \in U(\mathcal{H}_{S_2})$, then*

$$\begin{aligned} \text{Tr}_{ES_1} \left(V^{H_{S_2}}(U_2) V^{H_{S_1}}(U_1) \eta V^{H_{S_1}}(U_1)^\dagger V^{H_{S_2}}(U_2)^\dagger \right) \\ = \text{Tr}_E \left(V^{H_{S_2}}(U_2) \text{Tr}_{S_1}(\eta) V^{H_{S_2}}(U_2)^\dagger \right), \end{aligned} \quad (\text{A35})$$

for all $\eta \in \mathcal{S}(\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2} \otimes \mathcal{H}_E)$.

Strictly speaking we should rather write $[V^{H_{S_2}}(U_2) \otimes \hat{1}_{S_1}][\hat{1}_{S_2} \otimes V^{H_{S_1}}(U_1)]$ than $V^{H_{S_2}}(U_2)V^{H_{S_1}}(U_1)$ in the above proposition.

Proof. By Lemma 6 and cyclic permutation under the (partial) trace Tr_{ES_1} it follows that

$$\begin{aligned} & \text{Tr}_{ES_1} \left(V^{H_{S_2}}(U_2) V^{H_{S_1}}(U_1) \eta V^{H_{S_1}}(U_1)^\dagger V^{H_{S_2}}(U_2)^\dagger \right) \\ & \text{Tr}_{ES_1} \left(V^{H_{S_1}}(U_1) V^{H_{S_2}}(U_2) \eta V^{H_{S_2}}(U_2)^\dagger V^{H_{S_1}}(U_1)^\dagger \right) \\ & = \text{Tr}_{ES_1} \left([\hat{1}_{S_2} \otimes V^{H_{S_1}}(U_1)^\dagger V^{H_{S_1}}(U_1)] V^{H_{S_2}}(U_2) \eta V^{H_{S_2}}(U_2)^\dagger \right) \end{aligned} \quad (\text{A36})$$

According to Lemma 2, V preserves Hermitian conjugate, operator multiplication, and the identity $V^{H_{S_1}}(U_1)^\dagger V^{H_{S_1}}(U_1) = \hat{1}_{S_1}$. From this we can conclude equation (A35). \square

6. An example

Here we consider a simple illustration of the catalytic implementation of channels. Consider a collection of non-interacting two-level systems. We let $|\psi_0\rangle, |\psi_1\rangle$ denote the eigenvectors with corresponding eigenvalues $z_0 = 0, z_1 = 1$ of each system Hamiltonian. Suppose that all two-level systems initially are in the ground states $|\psi_0\rangle$, then the channels (A10) and (A12) take the form

$$\begin{aligned} \Phi_{\sigma,U}(|\psi_0\rangle\langle\psi_0|) &= |U_{00}|^2 |\psi_0\rangle\langle\psi_0| + |U_{10}|^2 |\psi_1\rangle\langle\psi_1| \\ &\quad + U_{10}^* U_{00} \text{Tr}(\Delta\sigma) |\psi_0\rangle\langle\psi_1| \\ &\quad + U_{10} U_{00}^* \text{Tr}(\Delta^\dagger\sigma) |\psi_1\rangle\langle\psi_0|, \\ \Lambda_{|\psi_0\rangle\langle\psi_0|,U}(\sigma) &= |U_{00}|^2 \sigma + |U_{10}|^2 \Delta^\dagger \sigma \Delta, \end{aligned}$$

where $U_{nn'} = \langle\psi_n|U|\psi_{n'}\rangle$.

Suppose we ideally would wish to put all the two-level systems into the state $|\phi\rangle = (|\psi_0\rangle - i|\psi_1\rangle)/\sqrt{2}$, where we sequentially use one and the same reservoir, initially in state $\sigma^{(0)}$. As a measure of the quality of this preparation we take the (square of the) fidelity $F_U := \langle\phi|\Phi_{\sigma^{(0)},U}(|\psi_0\rangle\langle\psi_0|)|\phi\rangle$, and optimize over all unitary U . An optimizing unitary is obtained by $U_{10} = 1/\sqrt{2}$ and $U_{00} = -i \exp[-i \arg(\text{Tr}(\Delta\sigma^{(0)}))]$, and leads to the optimal value $F_{\text{opt}}^{(0)} = \frac{1}{2} + \frac{1}{2} |\text{Tr}(\Delta\sigma^{(0)})|$. With this optimizing unitary, the next state of the reservoir is $\sigma^{(1)} := \frac{1}{2}\sigma^{(0)} + \frac{1}{2}\Delta^\dagger\sigma^{(0)}\Delta$. As seen (and as we already know) $\text{Tr}(\Delta\sigma^{(1)}) = \text{Tr}(\Delta\sigma^{(0)})$. Hence, in the second step we can use the same optimizing unitary, resulting in the same optimum $F_{\text{opt}}^{(1)} = F_{\text{opt}}^{(0)}$. Clearly, this process can be repeated indefinitely, with a constant sequence of fidelities $F_{\text{opt}}^{(k)} = \frac{1}{2} + \frac{1}{2} |\text{Tr}(\Delta\sigma^{(0)})|$ for all k . However, the state of the reservoir changes, and at the k th step it is

$$\sigma^{(k)} = \frac{1}{2^k} \sum_{l=0}^k \binom{k}{l} \Delta^{\dagger l} \sigma^{(0)} \Delta^l. \quad (\text{A37})$$

This provides an example of the spreading of the state mentioned several times in the main text. As a special case one can choose $\sigma^{(0)} := |\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|$, for $|\eta_{L,l_0}\rangle = \sum_{l=0}^{L-1} |l + l_0\rangle/\sqrt{L}$. In this case we obtain $F_{\text{opt}}^{(k)} = 1 - 1/(2L)$.

Appendix B: Half-infinite ladder (harmonic oscillator)

As mentioned earlier, a problem with the doubly-infinite energy ladder is that it is somewhat unphysical, as it has no ground state energy. However, by ‘cutting away’ the lower half of the spectrum we get the spectrum of the harmonic oscillator. Here we show that we can, with some modifications, reconstruct the catalytic properties of the doubly-infinite energy ladder in this harmonic oscillator model.

1. The half-infinite ladder model

Here we use an harmonic oscillator as model (or any Hamiltonian that is iso-spectral to the harmonic oscillator).

$$H_E^+ := s \sum_{j=0}^{+\infty} j |j\rangle\langle j|. \quad (\text{B1})$$

As before, we define an N -level system with a Hamiltonian $H_S \in H_s(\mathcal{H}_S)$, with eigenvalues $h_n = sz_n$. We also let

$$z_{\min} := \min_{n=1,\dots,N} z_n, \quad z_{\max} := \max_{n=1,\dots,N} z_n. \quad (\text{B2})$$

The projectors onto the eigenspaces of $H_S + H_E^+$ can be written

$$P_+^{(l)} := \sum_{n=1,\dots,N: l \geq z_n} |\psi_n\rangle\langle\psi_n| \otimes |l - z_n\rangle\langle l - z_n|, \quad (\text{B3})$$

for all $l \geq z_{\min}$. (It is maybe slightly annoying that we let the index l start at z_{\min} rather than at zero, but we do this for the sake of compactness of formulas and coordination with the previous section.) Note that

$$P_+^{(l)} = \sum_{n=1,\dots,N} |\psi_n\rangle\langle\psi_n| \otimes |l - z_n\rangle\langle l - z_n|, \quad \forall l \geq z_{\max}. \quad (\text{B4})$$

In other words, $P_+^{(l)} = P^{(l)}$ for $l \geq z_{\max}$. (We will in this section often compare operators on the spaces \mathcal{H}_E^+ and \mathcal{H}_E , where the former can be regarded as a subspace of the latter. Without further comments we do in these cases assume that the operators on \mathcal{H}_E^+ are extended trivially as zero operators to the orthogonal complement of \mathcal{H}_E^+ in \mathcal{H}_E , i.e., they act trivially on the negative half-ladder.)

For every $U \in U(\mathcal{H}_S)$ we define the unitary operator

$$V_+(U) := \sum_{l \geq z_{\max}} V_l(U) + \sum_{l=z_{\min}}^{z_{\max}-1} X_l \quad (\text{B5})$$

$$X_l X_l^\dagger = X_l^\dagger X_l = P_+^{(l)},$$

where

$$V_l(Q) := \sum_{n, n'=1}^N |\psi_n\rangle \langle \psi_n| Q |\psi_{n'}\rangle \langle \psi_{n'}| \otimes |l - z_n\rangle \langle l - z_{n'}|, \quad l \geq z_{\max} \quad (\text{B6})$$

is as in equation (A7). The unitarity of $V_+(U)$ follows from the fact that $V_l(U) V_l(U)^\dagger = V_l(U)^\dagger V_l(U) = P_+^{(l)}$. We will not pay any particular attention to the choice of the operators X_l , as we will focus on states outside the supports of these operators. As seen, we have, by the very construction, made sure that $V_+(U)$ in (A7) and $V(U)$ in (B5) act identically for states with support sufficiently far away from the ground state. The rest of this section is devoted to formalize the simple idea that if we start of the energy reservoir in a state that is sufficiently far from the ground state, we can maintain it so by suitable arrangements, and by doing this we retrieve all the relevant properties of the doubly-infinite ladder model.

Analogous to Δ we define the operator

$$\Delta_+ := \sum_{j=0}^{+\infty} |j+1\rangle \langle j|. \quad (\text{B7})$$

Note that Δ_+ is not unitary, but rather a partial isometry that takes \mathcal{H}_E^+ into the subspace $\text{Sp}\{|j\rangle\}_{j \geq 1}$. (This is of course possible only because \mathcal{H}_E^+ is infinite-dimensional.)

2. Sufficiently far from the ground state, the two models are equivalent

Analogous to the channel $\Phi_{\sigma, U}$ in Section A 2 we define

$$\Phi_{\sigma, U}^+(\rho) := \text{Tr}_E V_+(U) \rho \otimes \sigma V_+(U)^\dagger. \quad (\text{B8})$$

Similarly, we denote the set of channels that can be reached on S , using σ as a resource, by

$$\mathcal{C}_+^{S, H_S}(\sigma) := \left\{ \Phi_{\sigma, U}^+ : U \in U(\mathcal{H}_S) \right\}. \quad (\text{B9})$$

The effect of $V_+(U)$ on the energy reservoir is expressed via the channel

$$\Lambda_{\rho, U}^+(\sigma) := \text{Tr}_S V_+(U) \rho \otimes \sigma V_+(U)^\dagger. \quad (\text{B10})$$

Define the family of projectors

$$P_{\geq m} := \sum_{j \geq m} |j\rangle \langle j|, \quad m \geq 0. \quad (\text{B11})$$

Lemma 7. *Let $s > 0$, \mathcal{H}_S finite-dimensional, and $H_S \in H_s(\mathcal{H}_S)$ with maximal eigenvalue sz_{\max} and minimal eigenvalue sz_{\min} . Let $\sigma \in \mathcal{S}(\mathcal{H}_E)$ be such that $P_{\geq z_{\max}-z_{\min}} \sigma P_{\geq z_{\max}-z_{\min}} = \sigma$, then*

$$\begin{aligned} \Phi_{\sigma, U}^+ &= \Phi_{\sigma, U}, \\ \Lambda_{\rho, U}^+(\sigma) &= \Lambda_{\rho, U}(\sigma), \end{aligned} \quad (\text{B12})$$

for all $U \in U(\mathcal{H}_S)$.

Proof. By comparing with the definition of $P_+^{(l)}$ in (B3) one can see that

$$P_+^{(l)} [\hat{1}_S \otimes P_{\geq z_{\max}-z_{\min}}] = 0, \quad z_{\min} \leq l < z_{\max}. \quad (\text{B13})$$

Analogously

$$P^{(l)} [\hat{1}_S \otimes P_{\geq z_{\max}-z_{\min}}] = 0, \quad l < z_{\max}. \quad (\text{B14})$$

Note that $P_+^{(l)} = P^{(l)}$ and $V_l(U) P^{(l)} = V_l(U)$ for $l \geq z_{\max}$, and $X_l P_+^{(l)} = X_l$. By comparing (A5), (A7), and (B5), it follows that

$$V_+(U) [\hat{1}_S \otimes P_{\geq z_{\max}-z_{\min}}] = V(U) [\hat{1}_S \otimes P_{\geq z_{\max}-z_{\min}}]. \quad (\text{B15})$$

The statements of the lemma follows directly from this. \square

Lemma 8. *Let $s > 0$, \mathcal{H}_S finite-dimensional, and $H_S \in H_s(\mathcal{H}_S)$ with maximal eigenvalue sz_{\max} and minimal eigenvalue sz_{\min} . For $m \geq z_{\max} - z_{\min}$*

$$\begin{aligned} V_+(U) [\hat{1}_S \otimes P_{\geq m}] &= [\hat{1}_S \otimes P_{\geq m-z_{\max}+z_{\min}}] V_+(U) [\hat{1}_S \otimes P_{\geq m}]. \end{aligned} \quad (\text{B16})$$

Hence, if $\sigma \in \mathcal{S}(\mathcal{H}_E)$ is such that $P_{\geq m} \sigma P_{\geq m} = \sigma$, then

$$\begin{aligned} P_{\geq m-z_{\max}+z_{\min}} \Lambda_{\rho, U}^+(\sigma) P_{\geq m-z_{\max}+z_{\min}} &= \Lambda_{\rho, U}^+(\sigma), \\ \forall \rho \in \mathcal{S}(\mathcal{H}_S), \quad \forall U \in U(\mathcal{H}_S). \end{aligned} \quad (\text{B17})$$

Proof. Since $m \geq z_{\max} - z_{\min}$ it follows by equation (B13) that

$$\begin{aligned} V_+(U) [\hat{1}_S \otimes P_{\geq m}] &= \sum_{l \geq z_{\max}} \sum_{n, n'=1}^N |\psi_n\rangle \langle \psi_n| Q |\psi_{n'}\rangle \langle \psi_{n'}| \otimes |l - z_n\rangle \langle l - z_{n'}| P_{\geq m}. \end{aligned} \quad (\text{B18})$$

Let $|x\rangle$ be an energy eigenstate of the reservoir. For $\langle x | V_+(U) [\hat{1}_S \otimes P_{\geq m}] \neq 0$, a necessary condition is that $x = l - z_n$ and $l - z_{n'} \geq m$ for some l, n, n' . We can thus write $x + z_{\max} \geq x + z_n = l \geq m + z_{n'} \geq m + z_{\min}$. Hence, $\langle x | V_+(U) [\hat{1}_S \otimes P_{\geq m}] = 0$ for all $x < m - z_{\max} + z_{\min}$. It follows that $[\hat{1}_S \otimes (\hat{1}_E - P_{\geq m-z_{\max}+z_{\min}})] V_+(U) [\hat{1}_S \otimes P_{\geq m}] = 0$. This proves the lemma. \square

The following Proposition tells us that a reservoir with a state that is L times removed from the ground state with respect to the maximal possible energy change D , can be used L times before there is any noticeable difference in the set of channels that can be reached by using the reservoir. We refer to this as a quasi-catalytic property. Here $\lambda_{\max}(H)$ and $\lambda_{\min}(H)$ denote the maximal and minimal eigenvalues of the Hermitian operator H .

Proposition 4 (Quasi-Catalytic states). *Let $S > 0$, $D \in \mathbb{N}$, and let $\mathcal{H}_{S_1}, \dots, \mathcal{H}_{S_L}$ be finite-dimensional, with $H_{S_l} \in H_S(\mathcal{H}_{S_l})$, such that $\lambda_{\max}(H_{S_l}) - \lambda_{\min}(H_{S_l}) \leq sD$ for $l = 1, \dots, L$. Suppose $\sigma^{(1)} \in \mathcal{S}(\mathcal{H}_E)$ is such that*

$$P_{\geq LD} \sigma^{(1)} P_{\geq LD} = \sigma^{(1)}. \quad (\text{B19})$$

Define the sequence $\sigma^{(l+1)} = \Lambda_{\rho_l, U_l}^+(\sigma^{(l)})$, for $l = 1, \dots, L-1$, for $\rho_l \in \mathcal{S}(\mathcal{H}_{S_l})$ and $U_l \in U(\mathcal{H}_{S_l})$. Then

$$\mathcal{C}_+^{S_l, H_{S_l}}(\sigma^{(l)}) = \mathcal{C}_+^{S_l, H_{S_l}}(\sigma^{(1)}), \quad l = 1, \dots, L. \quad (\text{B20})$$

Proof. By a repeated use of Lemma 8, it follows that $P_{\geq (L-l+1)D} \sigma^{(l)} P_{\geq (L-l+1)D} = \sigma^{(l)}$. Thus, by Lemma 7 we know that $\Lambda_{\rho_l, U_l}^+(\sigma^{(l)}) = \Lambda_{\rho_l, U_l}(\sigma^{(l)})$. By Lemma 7 we can also conclude that $\mathcal{C}_+^{S_l, H_{S_l}}(\sigma^{(l)}) = \mathcal{C}^{S_l, H_{S_l}}(\sigma^{(l)})$. Furthermore, by Proposition 2 $\mathcal{C}^{S_l, H_{S_l}}(\Lambda_{\rho_{l-1}, U_{l-1}}(\sigma^{(l-1)})) = \mathcal{C}^{S_l, H_{S_l}}(\sigma^{(l-1)})$. By combining the above observations, we see that

$$\begin{aligned} \mathcal{C}_+^{S_l, H_{S_l}}(\sigma^{(l)}) &= \mathcal{C}_+^{S_l, H_{S_l}}(\Lambda_{\rho_{l-1}, U_{l-1}}^+(\sigma^{(l-1)})) \\ &= \mathcal{C}^{S_l, H_{S_l}}(\Lambda_{\rho_{l-1}, U_{l-1}}^+(\sigma^{(l-1)})) \\ &= \mathcal{C}^{S_l, H_{S_l}}(\Lambda_{\rho_{l-1}, U_{l-1}}(\sigma^{(l-1)})) \\ &= \mathcal{C}^{S_l, H_{S_l}}(\sigma^{(l-1)}) \\ &= \mathcal{C}_+^{S_l, H_{S_l}}(\sigma^{(l-1)}). \end{aligned} \quad (\text{B21})$$

The above reasoning can now be iterated, which yields the statement of the proposition. \square

3. Regenerative cycles

Here we construct a protocol to maintain the set of reachable channels \mathcal{C}_+ . Basically, what we do is to keep the state of the energy reservoir sufficiently far away from the ground state by pumping energy into it.

The following lemma tells us that we can implement the rigid translation $\sigma \mapsto \Delta_+^m \sigma \Delta_+^m$ within our model, if we have access to an ancillary system in a pure excited state. By the observation that $P_+^{(l)} |a_{\max}\rangle |j\rangle = 0$ for $z_{\max} > l \geq z_{\min}$ and all $j \geq 0$ we can directly use (B5) and (B6) to prove the following Lemma.

Lemma 9. *Let $H_A \in H_s(\mathcal{H}_A)$ have the largest eigenvalue sz_{\max} with corresponding eigenstate $|a_{\max}\rangle$, and smallest eigenvalue sz_{\min} with corresponding eigenstate*

$|a_{\min}\rangle$. Assume $z_{\max} \neq z_{\min}$. Let $U \in U(\mathcal{H}_A)$ be such that

$$U|a_{\max}\rangle = |a_{\min}\rangle. \quad (\text{B22})$$

Then

$$\Lambda_{|a_{\max}\rangle\langle a_{\max}|, U_A}^+(\sigma) = \Delta_+^{z_{\max}-z_{\min}} \sigma \Delta_+^{\dagger z_{\max}-z_{\min}}, \quad (\text{B23})$$

for all $\sigma \in \mathcal{S}(\mathcal{H}_E^+)$ where Λ^+ is as defined in equation (B10).

One may note that in this implementation there is no need to avoid any ‘border zone’.

In essence, the following proposition tells us that if the state of the energy reservoir has a support that is sufficiently far from the ground state, then we can maintain its power to induce channels on S (as characterized by the set \mathcal{C}_+^{S, H_S}) merely by feeding energy into it between each use.

Proposition 5 (Protocol for a catalytic cycle). *Let $s > 0$, $D \in \mathbb{N}$, $\mathcal{H}_{S_1}, \mathcal{H}_{S_2}$ be finite-dimensional, and $H_{S_1} \in H_s(\mathcal{H}_{S_1})$, $H_{S_2} \in H_s(\mathcal{H}_{S_2})$, where $\lambda_{\max}(H_{S_1}) - \lambda_{\min}(H_{S_1}) \leq sD$ and $\lambda_{\max}(H_{S_2}) - \lambda_{\min}(H_{S_2}) \leq sD$, with $D \in \mathbb{N}$. Let $\sigma_{\text{in}} \in \mathcal{S}(\mathcal{H}_E^+)$ be such that*

$$P_{\geq 2D} \sigma_{\text{in}} P_{\geq 2D} = \sigma_{\text{in}}. \quad (\text{B24})$$

For an arbitrary $\rho_1 \in \mathcal{S}(\mathcal{H}_{S_1})$ and an arbitrary $U_1 \in U(\mathcal{H}_{S_1})$, let

$$\bar{\sigma} := \Lambda_{\rho_1, U_1}^+(\sigma_{\text{in}}), \quad (\text{B25})$$

where the channel Λ_{ρ_1, U_1}^+ is as defined in equation (B10). Let \mathcal{H}_A be two-dimensional, with $H_A \in H_s(\mathcal{H}_A)$ such that H_A has the largest eigenvalue sD with corresponding eigenvector $|a_1\rangle$, and lowest eigenvalue 0 with corresponding eigenvector $|a_0\rangle$. Let $U_A = |a_1\rangle\langle a_0| + |a_0\rangle\langle a_1|$, and define

$$\sigma_{\text{out}} := \Lambda_{|a_1\rangle\langle a_1|, U_A}^+(\bar{\sigma}), \quad (\text{B26})$$

where $\Lambda_{|a_1\rangle\langle a_1|, U_A}^+$ is as defined in equation (B10). Then

$$\begin{aligned} \mathcal{C}_+^{S_2, H_{S_2}}(\sigma_{\text{out}}) &= \mathcal{C}^{S_2, H_{S_2}}(\sigma_{\text{out}}) \\ &= \mathcal{C}^{S_2, H_{S_2}}(\sigma_{\text{in}}) = \mathcal{C}_+^{S_2, H_{S_2}}(\sigma_{\text{in}}) \end{aligned} \quad (\text{B27})$$

and

$$P_{\geq 2D} \sigma_{\text{out}} P_{\geq 2D} = \sigma_{\text{out}}. \quad (\text{B28})$$

Proof. Combining (B24) and Lemma 8 yields

$$P_{\geq D} \bar{\sigma} P_{\geq D} = \bar{\sigma}. \quad (\text{B29})$$

By further combining this with Lemma 9 gives (B28). Due to the latter we can, by Lemma 7 conclude that

$$\mathcal{C}_+^{S_2, H_{S_2}}(\sigma_{\text{out}}) = \mathcal{C}^{S_2, H_{S_2}}(\sigma_{\text{out}}). \quad (\text{B30})$$

By (B29) and Lemma 8 it follows that $\sigma_{\text{out}} = \Lambda_{|a_1\rangle\langle a_1|, U_A}^+(\bar{\sigma}) = \Lambda_{|a_1\rangle\langle a_1|, U_A}(\bar{\sigma})$. Hence Proposition 2 tells us that

$$\mathcal{C}^{S_2, H_{S_2}}(\Lambda_{|a_1\rangle\langle a_1|, U_A}(\bar{\sigma})) = \mathcal{C}^{S_2, H_{S_2}}(\bar{\sigma}) \quad (\text{B31})$$

Due to (B24) and Lemma 8 it follows that $\Lambda_{\rho_1, U_1}^+(\sigma_{\text{in}}) = \Lambda_{\rho_1, U_1}(\sigma_{\text{in}})$. By Proposition 2

$$\mathcal{C}^{S_2, H_{S_2}}(\Lambda_{\rho_1, U_1}(\sigma_{\text{in}})) = \mathcal{C}^{S_2, H_{S_2}}(\sigma_{\text{in}}). \quad (\text{B32})$$

Furthermore, due to (B24) it follows by Lemma 7 that

$$\mathcal{C}^{S_2, H_{S_2}}(\sigma_{\text{in}}) = \mathcal{C}_+^{S_2, H_{S_2}}(\sigma_{\text{in}}). \quad (\text{B33})$$

By combining the above chain of equalities, we obtain the proposition. \square

Appendix C: More general models

In the previous sections we demonstrated that there in principle exist systems where coherence can be made catalytic. Here we briefly touch upon the question to what extent these phenomena survive in more general systems.

1. General energy preserving unitary operations

Here we consider Hamiltonians of the half-infinite ladder model, but we relax the restriction on the dynamics, and only demand that it preserves the total energy.

One can realize that a unitary operator V on the system and half-infinite ladder is block-diagonal with respect to the energy eigenspaces, i.e., $V = \sum_{l \geq z_{\min}} P_+^{(l)} V P_+^{(l)}$, if and only if

$$\begin{aligned} V = & \sum_{l \geq z_{\max}} \sum_{n, n'=1}^N V_{n, n'}^{(l)} |\psi_n\rangle\langle\psi_{n'}| \otimes |l - z_n\rangle\langle l - z_{n'}| \\ & + \sum_{l=z_{\min}}^{z_{\max}-1} X_l, \quad X_l X_l^\dagger = X_l^\dagger X_l = P_+^{(l)}, \end{aligned} \quad (\text{C1})$$

where each $V^{(j)} = [V_{n, n'}^{(j)}]_{n, n'=1}^N$ is a unitary matrix.

Recall that the harmonic oscillator model in Section B has the feature that it, sufficiently far from the ground state, essentially implements copies of a unitary interaction. We thus regain the model in Section B if we for all l choose $V^{(l)} = V$ for a fixed unitary matrix V . This ‘uniformity’ over the energy ladder appears to be closely connected to the catalytic properties of these models. It thus seems reasonable to suspect that in models that have some type of approximate uniformity, one should be able to find some approximate version of the type of behaviour that we have seen for the previous models.

Suppose that the set of unitary matrices $V^{(l_0)}, \dots, V^{(l_0+L)}$ would to become more ‘similar’

to each other as l_0 increases (for a fixed L). One way in which this could happen if $V^{(l)}$ would converge to a specific limit matrix as $l \rightarrow +\infty$. However, this is not the only possibility. For example, let H be an $N \times N$ Hermitian matrix, and define $V^{(l)} := \exp(-if(l)H)$, where $f : \mathbb{N} \rightarrow \mathbb{N}$ is such that $f(l+L) - f(l) \rightarrow 0$ as $l \rightarrow +\infty$. [An example of this is $f(x) = \sqrt{x}$, for which $\sqrt{x+L} - \sqrt{x} \sim L/(2\sqrt{x})$.] In this case $(V^{(l_0)}, \dots, V^{(l_0+L)}) \sim_{l_0 \rightarrow \infty} (V^{(l_0)}, \dots, V^{(l_0)})$.

Hence, for a limited range of levels, the transformations become asymptotically uniform (in spite the fact that they do not converge to a limiting operator). In Section C 2 we show that the Jaynes-Cummings model have this asymptotic uniformity.

As a final remark we note that in Section B we could maintain the capacity of the coherence by injecting energy into the reservoir. Furthermore, we did this solely ‘within’ the model, i.e., we used the same class of unitaries $V_+(U)$ for the pumping, as for the operations on system S . It is a relevant question for which classes of these more general unitary operations that it is possible (at least approximately). Unfortunately we do not provide an answer to this question here.

2. The Jaynes-Cummings model

Here we numerically analyze the Jaynes-Cummings model [1, 2] of a two-level system in resonant interaction with an electromagnetic field mode. As we shall see, numerical tests suggest that the higher the energy of a (suitably chosen) initial state of the reservoir, the longer its coherence survives sequential interactions with a collection of two-level systems.

To this end, we consider the Jaynes-Cummings model [1, 2] for the interaction between a two-level system $H_S = h_1|\psi_0\rangle\langle\psi_0| + h_2|\psi_1\rangle\langle\psi_1|$ and an electromagnetic field mode $H_E = \hbar\omega a^\dagger a$, with the interaction Hamiltonian

$$H_{\text{JC}} = g\sigma_+ \otimes a + g\sigma_- \otimes a^\dagger, \quad (\text{C2})$$

where a, a^\dagger are the standard annihilation and creation operators $[a, a^\dagger] = \hat{1}_E$, $a = \sum_{l=1}^\infty \sqrt{l} |l-1\rangle\langle l|$ and $\sigma_+ = |\psi_1\rangle\langle\psi_0|$, $\sigma_- = |\psi_0\rangle\langle\psi_1|$.

We furthermore assume resonant conditions (i.e., the ‘matching’ of the energy levels), so that $h_1 - h_0 = \hbar\omega$, and choose $h_1 = \hbar\omega z_1$, $h_0 = \hbar\omega z_0$, with $z_0 = 0, z_1 = 1$. By standard textbook calculations, where we introduce the unit-free evolution parameter $x := tg/\hbar$ for the evolution time t , the evolution induced by H_{JC} can be written

$$\begin{aligned} e^{-ixH_{\text{JC}}} = & \sum_{l=1}^\infty \sum_{n, n'=0}^1 V_{n, n'}^{(l)} |\psi_n\rangle\langle\psi_{n'}| \otimes |l - n\rangle\langle l - n'| \\ & + |\psi_0\rangle\langle\psi_0| \otimes |0\rangle\langle 0|, \\ V^{(l)} := & \begin{bmatrix} \cos(x\sqrt{l}) & -i \sin(x\sqrt{l}) \\ -i \sin(x\sqrt{l}) & \cos(x\sqrt{l}) \end{bmatrix} \end{aligned}$$

where we can write

$$V^{(l)} = e^{-ix\sqrt{l}H}, \quad H := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (\text{C3})$$

Hence, the dynamics in the JC-model provides an example of the type of asymptotically uniform unitary operators discussed in the previous section. The evolution on the reservoir is given by

$$\begin{aligned} \Lambda_{x,|\psi_0\rangle\langle\psi_0|}(\sigma) &:= \text{Tr}_S(e^{-ixH}|\psi_0\rangle\langle\psi_0| \otimes \sigma e^{ixH}) \\ &= A_x \sigma A_x^\dagger + B_x \sigma B_x^\dagger \\ A_x &:= \sum_{l=0}^{\infty} \cos(x\sqrt{l})|l\rangle\langle l| \\ B_x &:= \sum_{l=1}^{\infty} \sin(x\sqrt{l})|l-1\rangle\langle l|. \end{aligned}$$

The sequential preparation procedure results in a evolution process on the reservoir, which corresponds to an iteration of the channel $\Lambda_{x,|\psi_0\rangle\langle\psi_0|}$. The nature of this dynamics is depicted in Figure 4.

Here we attempt to mimic the sequential preparation procedure that we investigated in Section A 6. Again we thus use the (square of) the fidelity to measure how well the superposition $|\phi\rangle := (|\psi_0\rangle - i|\psi_1\rangle)/\sqrt{2}$ can be prepared from the ground state $|\psi_0\rangle$, by using the state σ on the reservoir as a resource. More precisely, for a given x and σ , our measure of ‘quality’ is

$$\begin{aligned} Q_x(\sigma) &:= \langle\phi|\Phi_{x,\sigma}(|\psi_0\rangle\langle\psi_1|)|\phi\rangle \\ &= \frac{1}{2} + \sum_{l=0}^{\infty} \sin(x\sqrt{l+1}) \cos(x\sqrt{l}) \text{Re}\langle l|\sigma|l+1\rangle \end{aligned} \quad (\text{C4})$$

Analogously to what we did in Section A 6 (where we optimized over all U) we should in principle optimize $Q_x(\sigma)$ over all $x \geq 0$. However, the dynamics of the JC-model can unfortunately be rather irregular (see e.g. [2, 40]) why such an optimization may not be feasible. Instead we opt for a simpler (but possibly suboptimal) procedure, where we keep the value of x fixed (corresponding to one fixed unitary). For the sake of simplicity we choose $x = \pi/4$. For an initial state $\sigma^{(0)}$ we construct the sequence $\sigma^{(k+1)} = \Lambda_{\pi/4,|\psi_0\rangle\langle\psi_0|}(\sigma^{(k)})$ and the corresponding fidelities $Q_{\pi/4}(\sigma^{(k)})$. For a suitably chosen l_0 the initial state is $\sigma^{(0)} := |\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|$. As we have seen, the evolution is asymptotically uniform. Hence, if L is small compared to l_0 , one can make the approximation

$$Q_{\pi/4}(\sigma^{(0)}) \approx \frac{1}{2} + \frac{1}{2}(1 - \frac{1}{L}) \sin(\pi\sqrt{l_0}/2). \quad (\text{C5})$$

and thus obtain the (approximate) maxima $1 - 1/(2L)$ at

$$l_0(m) := (4m+1)^2, \quad m = 1, 2, \dots \quad (\text{C6})$$

Each graph in Figure 3 in the main text is obtain from an initial state that is a uniform superpositions stretching

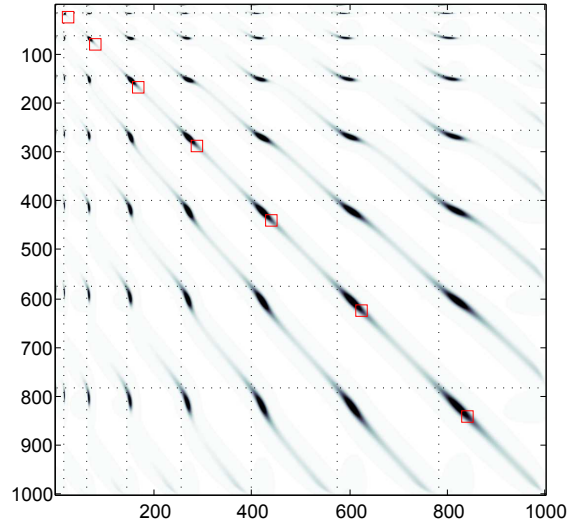


FIG. 4: **The state of the reservoir after iterated use.** For an initial state $\sigma^{(0)} = |\eta\rangle\langle\eta|$, with a very wide superposition $|\eta\rangle = \sum_{l=0}^{1000} |l\rangle/\sqrt{1000}$, the figure depicts $|\langle l|\sigma^{(100)}|l'\rangle|$ (black corresponds to high values, and white low) where $\sigma^{(100)}$ is the 100th iteration of $\sigma^{(k+1)} = \Lambda_{\pi/4,|\psi_0\rangle\langle\psi_0|}(\sigma^{(k)})$, for $x = \pi/4$. The dotted lines corresponds to the values of l (and l') for which $B_{\pi/4}|l\rangle = 0$ ($B_{\pi/4}|l'\rangle = 0$). Since B_x induces loss of quanta from the reservoir, this suggests that weight will tend to accumulate at the crossings of the dotted lines. This figure shows an intermediate step in this process. The squares show the points $l_0(m) = (4m+1)^2$ around which $Q_{\pi/4}$ takes a high value (for relatively narrow distributions). On these points we center the family of initial states that lead to the curves in Figure 3. (As opposed to the $\sigma^{(0)}$ that yields the present figure, those initial states only contains a narrow distribution of number states.)

over 50 number states, centered around $l_0(m)$. Hence, $\sigma_m^{(0)} = |\eta^{(m)}\rangle\langle\eta^{(m)}|$, where $|\eta^{(m)}\rangle = \sum_{l=-24}^{25} |l_0(m) + l\rangle$. Each graphs corresponds to the value $F_k^{(m)} := Q_{\pi/4}(\sigma_m^{(k)})$ on the y -axis versus the number of iterations k on the x -axis. The different graphs correspond to the different values of m , where in the rightmost part of the figure, the curves correspond to $m = 4, 6, 8, \dots, 28, 30$ counted from below. These graphs suggest that the ‘lifetime’ of the coherence can be made longer, merely by increasing the average energy of the state.

Since these graphs in Figure 3 are the result of an approximate optimization, one should maybe not get too surprised by the fact that one see an initial increase of the fidelity in Figure 3. As seen, the largest increase happens for the initial states with a relatively low number of quanta, which is where we indeed would expect the approximation to be the worst. The dotted line in Figure 3 corresponds to the fidelity we would have obtained for the doubly-infinite ladder-model, i.e., $1 - 1/(2L)$ (see Section A 6). Note that there is no reason to expect this

to be a bound for what can be achieved in the JC model. However, the gradual approach of the curves to this value suggests that the applied approximation is reasonable.

One may wonder why we here use $|\eta_{L,l_0}\rangle$ as initial states, rather than coherent states, $|\alpha\rangle := e^{-|\alpha|^2/2} \sum_{l=0}^{+\infty} \frac{\alpha^l}{(l!)^{1/2}} |l\rangle$ [24–26], which would be the both convenient and conventional choice for this type of model. For the coherent states we could create a sequence of increasing average energy by increasing $|\alpha|$. However, as we increase $|\alpha|$ we also increase the width of the superposition over the energy eigenstates. Hence, in some intuitive sense the degree of coherence increases as we increase $|\alpha|$, why it appears risky to use these states for the questions we investigate here. One should keep in mind though, that since we have not developed any operational measure of the degree of coherence, discussions of this kind is a bit shaky. (For the doubly-infinite and the half-infinite ladder models, we could side-step the issue of measures of coherence by focusing on the set of induced channels $\mathcal{C}(\sigma)$.)

Appendix D: Prelude. Expected work-extraction without an energy reservoir

Before we begin with the actual purpose of analyzing the role of coherence in expected work extraction by using explicit energy reservoirs (which we do in Section E) we first consider some simpler types of models that do not contain an explicit energy reservoir, and where one can show that the optimal expected work extraction (suitably defined) is given by the relative von Neumann entropy $kTD(\rho\|G(H_S))$, where $D(\rho\|\sigma) = \text{Tr}(\rho \ln \rho) - \text{Tr}(\rho \ln \sigma)$. The purpose of this exposition is partially to highlight the difference between using, and not using, an explicit energy reservoir, but also to gain some understanding for the ideas behind the (somewhat technically messy) derivations in Section E. On a technical level, the only things that we need from this section are Lemma 11 and Corollary 2.

In Section D1 we begin with a short ‘prelude to the prelude’, where we briefly recall the common approach to implement expected work extraction using time-dependent Hamiltonians. More precisely, the system traverses through a path of Hamiltonians. These paths of Hamiltonians will implicitly play a role in the subsequent derivations, as we, in some sense, will simulate such paths. In Section D2 we recall the concept of passivity, and the related setting for work extraction. In Section D3 we implement one version of the above mentioned ‘simulation’, and show how this leads to optimal expected work extraction.

1. Expected work extraction by varying the Hamiltonian

A common approach (that can take many different guises on the level of assumptions and technical details) is to define expected work extraction in terms of changes of the Hamiltonian of the system. (For a handful of examples, see e.g. [20, 36, 41–45], where [45] provides an introduction.) In this setting we are given quantum states ρ and Hamiltonian H , and are allowed changes of Hamiltonians $H \mapsto H'$. The expected work gain of such a step is defined as

$$W_{\text{yield}}(H_S \mapsto H'_S, \rho) := \text{Tr}(H_S \rho) - \text{Tr}(H'_S \rho). \quad (\text{D1})$$

The contact with the heat bath is modeled as a replacement map $\rho \mapsto \Phi_H(\rho) := G(H)$. (As we will come back to below, there are variations on this, which can implement ‘softer’ types of thermalizations.) A process is defined as a sequence of shifts of Hamiltonians, sandwiched by thermalizations, where the total work gain is defined as the sum of all the individual gains as in (D1). This model is, for expected work extraction, a quantum generalization of the discrete classical model used in [14]. (Unfortunately, this generalization does not make much sense as a generalization for ϵ -deterministic work extraction.) Very much analogous to [14], one can show that for a cyclic change of Hamiltonians, the optimal expected work extraction in this setting becomes $kTD(\rho\|G(H_S))$. To see this, one can use the relation $H = F(H) - kT \ln G(H)$, where $F(H) = -kT \ln Z(H)$, and $Z(H) = \text{Tr} e^{-\beta H}$, to show that the total work gain of an L -step cyclic process ($H^{(0)} = H^{(L)} = H_S$), with $\rho^{(0)} := \rho$ and $\rho^{(l+1)} = \Phi_{H^{(l+1)}}(\rho^{(l)})$ becomes

$$\begin{aligned} W_{\text{yield}} &= \sum_{l=0}^{L-1} [\text{Tr}(H^{(l)} \rho_l) - \text{Tr}(H^{(l+1)} \rho_l)] \\ &= \frac{1}{\beta} D(\rho\|G(H_S)) - \frac{1}{\beta} D(\rho_{L-1}\|G(H_S)) \\ &\quad - \frac{1}{\beta} \sum_{l=1}^{L-1} [D(\rho_{l-1}\|G(H^{(l)})) - D(\rho_l\|G(H^{(l)}))]. \end{aligned} \quad (\text{D2})$$

Using $\Phi_H(G(H)) = G(H)$, together with the fact that the von-Neumann relative entropy is contractive under channels, i.e., $D(\Gamma(\rho)\|\Gamma(\sigma)) \leq D(\rho\|\sigma)$, it follows that $D(\rho_l\|G(H^{(l)})) = D(\Phi_{H^{(l)}}(\rho_{l-1})\|\Phi_{H^{(l)}}(G(H^{(l)}))) \leq D(\rho_{l-1}\|G(H^{(l)}))$, and thus $W_{\text{yield}} \leq kTD(\rho\|G(H_S)) - kTD(\rho_{L-1}\|G(H_S))$. Due to the fact that $D(\cdot\|\cdot) \geq 0$, it follows that within this model, for any cyclic process, we get the general bound $W_{\text{yield}} \leq kTD(\rho\|G(H_S))$.

As a side-remark one may note that, for the above derivation, we do not need to use the replacement map $\rho \mapsto G(H)$. As seen, we could, instead of Φ_H , use *any* channel that has $G(H)$ as a fix-point, i.e., $\Phi(G(H)) = G(H)$. (This implies a considerable freedom in implementing the softer types of thermalizations mentioned

above. However, we shall not need this freedom in our derivations, but will stay put with the, rather brutal, complete thermalization implemented by the replacement map.)

The next question is if it is possible to saturate the bound $W_{\text{yield}} \leq kTD(\rho\|G(H_S))$ within this model. A (limit) process that does achieve the bound can be described as follows: Change the initial Hamiltonian H_S into H' , where the latter is such that $\rho = G(H')$. (Strictly speaking, this only works if ρ has full support. This can be treated via yet another limiting process, with a sequence of Gibbs states that converges to the state, akin to what was done in [14].) Next, we thermalize the system and perform an alternating sequence of thermalizations and small changes of the Hamiltonians, keeping the eigenbasis fixed, until we reach a Hamiltonian H'' that is iso-spectral to H_S . Finally, we perform an alternating sequence of thermalizations and small unitary transformations, until H'' has been rotated back to H_S . The first of these steps has the expected work yield $W_{\text{yield}}^{(1)} = F(H_S) - F(H') + kTD(\rho\|G(H_S))$. The second step, does in the limit of an infinitely fine discretization, yield $W_{\text{yield}}^{(2)} = F(H') - F(H'')$. (Since we only have states that are diagonal in a fixed energy eigenbasis, the results for the ITR in [14] are applicable.) For the final step of the protocol, let U be a unitary that transforms an eigenbasis of H'' into an eigenbasis of H_S . We diagonalize $U = \sum_n e^{-i\phi_n} |\xi_n\rangle\langle\xi_n|$, for $0 \leq \phi_n < 2\pi$, and define the Hermitian operator $A := \sum_n \phi_n |\xi_n\rangle\langle\xi_n|$. We can thus define a sequence of Hamiltonians $H^{(l)} := e^{-iAl/L} H'' e^{iAl/L}$, for $l = 0, \dots, L$, such that $H^{(0)} = H''$ and $H^{(L)} = H_S$. Consider a process the sequentially changes $H^{(l)}$ to $H^{(l+1)}$, sandwiched by thermalizations. The work yield of this process is.

$$W_{\text{yield}}^{(3)} = \sum_{l=0}^{L-1} \text{Tr}[(H^{(l)} - H^{(l+1)})G(H^{(l)})] \quad (\text{D3})$$

$$= L\text{Tr}[(H'' - e^{-iA/L} H'' e^{iA/L})G(H'')].$$

By inserting the expansion $e^{-iA/L} H'' e^{iA/L} = H'' - i[A, H'']/L + O(1/L^2)$ into the above equation, one sees that $W_{\text{yield}}^{(3)} \rightarrow 0$ as $L \rightarrow \infty$. Hence, the total work yield $W_{\text{yield}} = W_{\text{yield}}^{(1)} + W_{\text{yield}}^{(2)} + W_{\text{yield}}^{(3)}$ can be made arbitrarily close to $kTD(\rho\|G(H_S))$.

2. Work extraction and passive states

The concept of passive states were introduced in the context of axiomatic characterizations of Gibbs states [46, 47]. See also recent results in [48] and [49] which study the role of entanglement in this type of setting.

We are given a quantum system S , with a Hamiltonian H_S , and a quantum state ρ . We are allowed to perform arbitrary unitary operations (no matter whether they mix energy levels or not). The work gain of a uni-

tary operation U is defined as

$$W_{\text{yield}}(U, \rho, H_S) := \text{Tr}(H_S \rho) - \text{Tr}(H_S U \rho U^\dagger). \quad (\text{D4})$$

In light of the results in the previous sections, one may suspect that this construction implicitly assumes the access to ideal coherence. To analyze this issue in a quantitative manner is precisely the purpose of Section E.

A state ρ is *passive* with respect to a given Hamiltonian H_S , if we cannot extract any energy from it in the sense of (D4), i.e., $W_{\text{yield}}(U, \rho, H_S) \leq 0$ for all unitary U [46, 47]. It turns out [46, 47] that a state is passive if and only if the state ρ and the Hamiltonian H_S has a common eigenbasis, and that in this basis the eigenvalues of the density operator are ordered non-increasingly with increasing energy eigenvalues (or equivalently if the eigenvalues of ρ and $G(H_S)$ co-decrease).

It is straightforward to calculate the optimal extraction (e.g., by use of Lemma 12)

$$W_{\text{yield}}^{\text{optimal}} := \sup_{U \in U(\mathcal{H}_S)} W_{\text{yield}}(U)$$

$$= \text{Tr}(H_S \rho_S) - \inf_{U \in U(\mathcal{H}_S)} \text{Tr}(H_S U \rho_S U^\dagger)$$

$$= \frac{1}{\beta} D(\rho_S \| G(H_S)) - \frac{1}{\beta} D(\lambda^\downarrow(\rho_S) \| \lambda^\downarrow(G(H_S))). \quad (\text{D5})$$

Here $\lambda_1^\downarrow(\rho) \geq \dots \geq \lambda_N^\downarrow(\rho)$ means the eigenvalues of the operator ρ_S ordered non-increasingly. As seen from (D5) we obtain the optimal work extraction precisely when we put the system in a passive state. One might wonder from where the β in (D5) comes from, as we have not yet introduced any heat bath. At this stage it is merely a mathematical identity. [All the β on the right hand side of (D5) cancel, for any non-zero β .]

Due to the fact that relative entropy is non-negative $D(\cdot\|\cdot) \geq 0$, the optimal value can never be larger than $kTD(\rho\|G(H_S))$. If we somehow could ‘arrange’ for the term $kTD(\lambda^\downarrow(\rho)\|\lambda^\downarrow(G(H_S)))$ to become zero, we would obtain $kTD(\rho\|G(H_S))$ as the optimum.

An important observation is that a combination of passive states is not necessarily passive [46, 47]. As an example, for a passive state ρ , the tensor product $\rho \otimes \rho$ may not necessarily be passive. For this we assume that the two subsystems have identical and non-interaction Hamiltonians, resulting in the total Hamiltonian $H_{S_1} \otimes \hat{1}_{S_2} + \hat{1}_{S_1} \otimes H_{S_2}$. We furthermore allow arbitrary unitary operations on the *joint* system $\mathcal{H}_{S_1} \otimes \mathcal{H}_{S_2}$. The Gibbs state $G(H)$ have the special property that it is completely passive [46, 47], meaning that it remains passive for arbitrary tensor products $G(H)^{\otimes n}$.

Combining copies of a system is not the only way in which we can change a passive state into a non-passive state. Another method is to add on a ancillary system A in a Gibbs state $G(H_A)$ for some Hamiltonian H_A . We furthermore assume that the ancillary system is non-interacting with S , such that the total Hamiltonian is $H_{SA} := H_S \otimes \hat{1}_A + \hat{1}_S \otimes H_A$. Substituting ‘ S ’ in (D5)

with SA with the new Hamiltonian H_{SA} , and the new initial state $\rho \otimes G(H_A)$ [where we let the β in (D5) be the same as the β in the Gibbs state $G(H_A)$ of the ancillary system] yields

$$\begin{aligned} W_{\text{yield}}^{\text{opt}} &= \text{Tr}((H_S + H_A)\rho \otimes G(H_A)) \\ &\quad - \inf_{U \in U(\mathcal{H}_S \otimes \mathcal{H}_A)} \text{Tr}((H_S + H_A)U\rho \otimes G(H_A)U^\dagger) \\ &= \frac{1}{\beta} D(\rho \| G(H_S)) \\ &\quad - \frac{1}{\beta} D(\lambda^\downarrow(\rho \otimes G(H_A)) \| \lambda^\downarrow(G(H_S) \otimes G(H_A))). \end{aligned} \quad (\text{D6})$$

Comparing (D6) with (D5) we see that the first term is unchanged [due to fact that $D(\rho \otimes G(H_A) \| G(H_S) \otimes G(H_A)) = D(\rho \| G(H_S))$] while the second term is altered. In general $D(\lambda^\downarrow(\rho \otimes G(H_A)) \| \lambda^\downarrow(G(H_S) \otimes G(H_A)))$ can be smaller than $D(\lambda^\downarrow(\rho) \| \lambda^\downarrow(G(H_S)))$, reflecting the fact that we can make a passive state non-passive by adding an ancillary Gibbs state. The purpose of the following section is to show that the former term can be made arbitrarily small by choosing the ancillary Hamiltonian H_A appropriately.

The justification for the use of these ancillary Gibbs states is that we in principle always can obtain them by equilibrating the desired ancillary system with respect to the given heat bath of temperature T , i.e., an ancillary Gibbs state is a ‘free resource’ [9, 10, 15]. More precisely, one can imagine that the ancilla system is put in contact with the heat bath. Next, we let it continuously equilibrate with the heat bath while we slowly de-connect it. (Since we here deal with questions about ideal thermodynamics, we do not worry about equilibration times.)

3. Simulation of time-dependent Hamiltonians using a sequence of subsystems

As is well known, relative entropy does not satisfy the triangle inequality. For the construction of our proofs, we make use of this ‘failure’ of the relative entropy. Lemma 11 shows that if we would regard the relative entropy as a (strange) distance measure, then the total sum of the ‘lengths’ of the pieces of a path can be made arbitrarily small merely by making the division finer. (This is very closely related to the ‘isothermal reversible paths’ as described in [14].)

Analogously to how we defined $G(H)$, $Z(H)$, $F(H)$, etc., for operators, we here analogously define them for real vectors $h = (h_1, \dots, h_N) \in \mathbb{R}^N$, by $G_n(h) := e^{-\beta h_n} / Z(H)$, $Z(h) := \sum_n e^{-\beta h_n}$, $F(h) := -\frac{1}{\beta} \ln Z(h)$.

Lemma 10. *Let $h : [0, 1] \rightarrow \mathbb{R}^N$ have component-wise continuous second derivatives, then*

$$\left| \frac{d^2}{dx^2} F(h(x)) \right| \leq \max_n |h_n''(x)| + 2\beta \max_n |h_n'(x)|^2. \quad (\text{D7})$$

By combining the above lemma with the Taylor expansions, $f(x) = f(x_0) + (x - x_0)f'(x_0) + r(x)$, with the error term in the Lagrange form $r(x) := \frac{(x - x_0)^2}{2} f''(\xi)$ for $\xi \in [\min(x, x_0), \max(x, x_0)]$ one can obtain the following lemma.

Lemma 11. *Let $N \in \mathbb{N}$ and $\beta > 0$ be given, and let $h : [0, 1] \rightarrow \mathbb{R}^N$ have component-wise continuous second derivatives. For each integer $K \geq 2$*

$$\begin{aligned} 0 &\leq \sum_{k=0}^{K-1} D\left(G\left(h\left(\frac{k}{K}\right)\right) \left\| G\left(h\left(\frac{k+1}{K}\right)\right)\right\| \\ &\leq \frac{1}{K} \left[\beta \max_{x \in [0, 1]} \max_n |h_n''(x)| + \beta^2 \max_{x \in [0, 1]} \max_n |h_n'(x)|^2 \right]. \end{aligned} \quad (\text{D8})$$

In the following we let $\mathbb{P}(N)$ denote the set of probability distributions over N symbols, and $\mathbb{P}_+(N)$ the subset of probability distributions with full support.

Lemma 12 (Corollary II.4.4 on p. 49 in [50]). *Let $a, b \in \mathbb{R}^M$, then*

$$\sum_{m=1}^M a_m^\downarrow b_m^\uparrow \leq \sum_{m=1}^M a_m b_m \leq \sum_{m=1}^M a_m^\uparrow b_m^\uparrow. \quad (\text{D9})$$

Any permutation π on the set $\{1, \dots, M\}$ induces an operation Π on $\mathbb{P}(M)$ by $\Pi_\pi(p)_m := p_{\pi(m)}$, i.e., we permute the elements in the index of p . A direct consequence of Lemma 12 is the following

Corollary 2. *Let $q \in \mathbb{P}(M)$ and $p \in \mathbb{P}_+(M)$, and π be any permutation on $\{1, \dots, M\}$, then*

$$D(q^\downarrow \| p^\downarrow) \leq D(\Pi_\pi(q) \| p). \quad (\text{D10})$$

Suppose that $M := N^L$, and suppose moreover that we choose to represent the elements of $\mathbb{P}(N^L)$ as L -dimensional tensors, i.e., as $(p_{n_1, \dots, n_L})_{(n_1, \dots, n_L)}$, where $(n_1, \dots, n_L) \in \{1, \dots, N\}^{\times L}$. We can obtain a special permutation on the index set $\{1, \dots, N\}^{\times L}$ by permuting the ‘index of the indices’, i.e., (n_1, \dots, n_L) is mapped to $(n_{\pi(1)}, \dots, n_{\pi(L)})$, where π is a permutation of $\{1, \dots, L\}$. This is of course merely a special case of a general permutation of $\{1, \dots, N\}^{\times L}$. This permutation of the index of indices induces a corresponding operation Π on $\mathbb{P}(N^L)$, by $\Pi((p_{n_1, \dots, n_L})_{n_1, \dots, n_L}) = (p_{n_{\pi(1)}, \dots, n_{\pi(L)}})_{n_1, \dots, n_L}$. As this is a special case of the more general permutation, it follows that Corollary 2 is applicable, which is important for the the proof of the following proposition.

Lemma 13. *Let N and $q, p \in \mathbb{P}_+(N)$ be given. For each integer $K \geq 2$ there exists a $h^K \in \mathbb{R}^{N^{K+1}}$ such that*

$$D\left((qG(h^K))^\downarrow \| (pG(h^K))^\downarrow\right) \leq \frac{1}{K} \max_n \left| \ln \frac{q_n}{p_n} \right|^2. \quad (\text{D11})$$

Proof. Since, $q, p \in \mathbb{P}_+(N)$, it follows that $\log q_n$ and $\log p_n$ are well defined and finite. Thus we can define

$h_n^i := -\frac{1}{\beta} \ln q_n$, $h_n^f := -\frac{1}{\beta} \ln p_n$. We connect these two points with the path $h : [0, 1] \rightarrow \mathbb{R}^N$ by $h(x) := (1-x)h^i + xh^f$. Lemma 11 is applicable, and thus delivers $h^{k:K} := h(\frac{k}{K})$ with $k = 0, 1, \dots, K$ such that

$$\sum_{k=0}^{K-1} D(G(h^{k:K}) \| G(h^{(k+1):K})) \leq \frac{1}{K} \max_n \left| \ln \frac{q_n}{p_n} \right|^2. \quad (\text{D12})$$

Define

$$h_{n_0, \dots, n_K}^K := h_{n_0}^{0:K} + \dots + h_{n_K}^{K:K}, \quad (n_0, \dots, n_K) \in \{1, \dots, N\}^{K+1}. \quad (\text{D13})$$

Hence $h^K \in \mathbb{R}^{N^{K+1}}$, and

$$G_{n_0, n_1, \dots, n_K}(h^K) = G_{n_0}(h^{0:K}) \dots G_{n_K}(h^{K:K}). \quad (\text{D14})$$

On $\mathbb{R}^{N^{K+2}}$ we define the operation Π by

$$(\Pi a)_{n_s, n_0, n_1, \dots, n_{K-1}, n_K} := a_{n_0, n_1, \dots, n_{K-1}, n_K, n_s}. \quad (\text{D15})$$

In other words, Π is defined via a cyclic permutation of the index of the indices of a . By Corollary 2 it follows that

$$\begin{aligned} & D((qG(h^K))^\downarrow \| (pG(h^K))^\downarrow) \\ & \leq D(\Pi(qG(h^K)) \| pG(h^K)) \\ & = D(G(h^{K:K})qG(h^{0:K}) \dots G(h^{K-1:K}) \| \\ & \quad pG(h^{0:K}) \dots G(h^{K-1:K})G(h^{K:K})) \quad (\text{D16}) \\ & = D(q \| G(h^{0:K})) + D(G(h^{K:K}) \| p) \\ & \quad + \sum_{k=0}^{K-1} D(G(h^{k:K}) \| G(h^{(k+1):K})). \end{aligned}$$

By construction $q = G(h^{0:K})$ and $p = G(h^{K:K})$, and by (D12) the statement of the lemma follows. \square

Proposition 6. Let $\rho \in \mathcal{S}_+(\mathcal{H}_S)$ and let H_S be a Hermitian operator on \mathcal{H}_S . Then there exists a sequence of ancillary Hilbert spaces \mathcal{H}_A^K and hermitian operators H_A^K such that

$$\begin{aligned} W_{\text{yield}}^{\text{opt}} &= \lim_{K \rightarrow +\infty} \left[\text{Tr}((H_S + H_A^K)\rho \otimes G(H_A^K)) \right. \\ & \quad \left. - \inf_{U \in U(\mathcal{H}_S \otimes \mathcal{H}_A^K)} \text{Tr}((H_S + H_A^K)U\rho \otimes G(H_A^K)U^\dagger) \right] \\ &= \frac{1}{\beta} D(\rho \| G(H_S)). \end{aligned} \quad (\text{D17})$$

Proof. Let q be the eigenvalues of ρ and p the eigenvalues of $G(H_S)$. Since $\rho \in \mathcal{S}_+(\mathcal{H}_S)$ it follows that $q \in \mathbb{P}_+(N)$. As p are the eigenvalues of a Gibbs state, it follows directly that $p \in \mathbb{P}_+(N)$. By a repeated use of Lemma 13 we can construct a sequence $(h^K)_{K \in \mathbb{N}}$,

$h^K \in \mathbb{R}^{N^{K+1}}$, where each h^K satisfies equation (D11) in Lemma 13. Let $\{|a_k^K\rangle\}_{k=1}^{N^{K+1}}$ be an orthonormal basis in an N^{K+1} -dimensional Hilbert space \mathcal{H}_A^K and define the Hermitian operator $H_A^K := \sum_{k=1}^{N^{K+1}} h_k^K |a_k^K\rangle\langle a_k^K|$. One can see that $D(\lambda^\downarrow(\rho \otimes G(H_A^K)) \| \lambda^\downarrow(G(H_S) \otimes G(H_A^K))) = D((qG(h^K))^\downarrow \| (pG(h^K))^\downarrow)$. Combining this observation with equations (D6) and (D11), we obtain $0 \leq kTD(\rho \| G(H_S)) - W_{\text{yield}}^{\text{opt}} \leq \frac{1}{K} \max_n \left| \ln \frac{q_n}{p_n} \right|^2$, which proves the proposition. \square

Appendix E: Expected work extraction with an explicit energy reservoir

In the context of passivity one defines (as we saw in Section D 2) the work gain as the decrease of internal energy of the system of interest under unitary operations. Here we introduce an explicit energy reservoir, and thus model the system into which the extracted energy is to be put. Consequently we define the expected work gain as the increase of expected energy in the reservoir. Throughout this section we will exclusively make use of the model introduced in Section A.

1. Expected work extraction in fixed systems

a. Operations on the system and energy reservoir

Definition 2. Let $s > 0$, let \mathcal{H}_S be a finite-dimensional Hilbert space, with $H_S \in \mathcal{H}_s(\mathcal{H}_S)$, and let $\sigma \in \mathcal{S}(\mathcal{H}_E)$. Let sz_k and $|\psi_k\rangle$ be eigenvalues and corresponding orthonormal eigenvectors of H_S . Define

$$\begin{aligned} R_\sigma^{H_S}(Q) &:= \sum_{n, n'=1}^N |\psi_n\rangle\langle\psi_n|Q|\psi_{n'}\rangle\langle\psi_{n'}| \text{Tr}(\Delta^{z_n - z_{n'}} \sigma), \\ &\quad \forall Q \in L(\mathcal{H}_S). \end{aligned} \quad (\text{E1})$$

The channel $R_\sigma^{H_S}$ should not be confused with $\Phi_{\sigma, U}^S$ introduced in Section A 2.

Lemma 14. For $s > 0$, $H_S \in \mathcal{H}_s(\mathcal{H}_S)$, and $\sigma \in \mathcal{S}(\mathcal{H}_E)$

$$R_\sigma^{H_S}(\hat{1}) = \hat{1}. \quad (\text{E2})$$

In other words, the channel is unital and can thus not decrease the von Neumann entropy.

So far in this investigations we have only dealt with expectation values of bounded operators, and have thus not needed to worry about the existence of these expectation values. However, in this section we need to consider expressions like $\text{Tr}(H_E^{(s)} \sigma)$. Since $H_E^{(s)}$ is an unbounded operator, we cannot guarantee that the expectation value is well defined for all elements $\sigma \in \mathcal{S}(\mathcal{H}_E)$. One way of dealing with this would be to restrict $\mathcal{S}(\mathcal{H}_E)$ to elements

for which the product $H_E^{(s)}\sigma$ is a trace class operator (Q is trace class if $\text{Tr}\sqrt{Q^\dagger Q} < +\infty$). In essence, we need a restriction such that $\langle n|\sigma|n'\rangle$ converges sufficiently fast to zero as $n, n' \rightarrow \pm\infty$. However, for the purposes of this investigation there appears to be no strong reason to consider this technical issue in detail, why we in the following merely let $\mathcal{S}^*(\mathcal{H}_E)$ denote some restriction of $\mathcal{S}(\mathcal{H}_E)$, where we (in one way or another) have made sure that all the relevant expectation values make sense. Note that all the operations we perform on the energy reservoir merely involve the interactions with a finite-dimensional space, why there is no reason to expect that these operations brings the state out of (a reasonably defined) $\mathcal{S}^*(\mathcal{H}_E)$.

Lemma 15. *Let $s > 0$. Let \mathcal{H}_S be a finite-dimensional Hilbert space, $H_S \in H_s(\mathcal{H}_S)$, $\rho \in \mathcal{S}(\mathcal{H}_S)$, and $\sigma \in \mathcal{S}^*(\mathcal{H}_E)$. Then*

$$\begin{aligned} & \text{Tr}\left([\hat{1}_S \otimes H_E^{(s)}]V(U)\rho \otimes \sigma V(U)^\dagger\right) \\ &= \text{Tr}(H_E^{(s)}\sigma) + \text{Tr}(H_S\rho) \\ & \quad - \text{Tr}\left(U^\dagger H_S U R_\sigma^{H_S}(\rho)\right), \end{aligned} \quad (\text{E3})$$

for all $U \in \mathbb{U}(\mathcal{H}_S)$ and all $\rho \in \mathcal{S}(\mathcal{H}_S)$, where V is defined as in Lemma 2, and $R_\sigma^{H_S}$ is as in Definition 2.

Proof. The first part of the proof is to rewrite $\text{Tr}([\hat{1}_S \otimes H_E^{(s)}]V(U)\rho \otimes \sigma V(U)^\dagger) = \text{Tr}([H_S \otimes \hat{1}_E + \hat{1}_S \otimes H_E^{(s)}]\rho \otimes \sigma) - \text{Tr}([H_S \otimes \hat{1}_E]V(U)\rho \otimes \sigma V(U)^\dagger)$, where we have used the fact that $V(U)$ is energy conserving. By using equation (A5) in Lemma 2, and use the fact that H_S is diagonal in the energy eigenbasis $\{|\psi_n\rangle\}_{n=1}^N$, one can see that $\text{Tr}([H_S \otimes \hat{1}_E]V(U)\rho \otimes \sigma V(U)^\dagger) = \sum_{k,l,l'} s z_{l'} \langle \psi_l | U^\dagger | \psi_{l'} \rangle \langle \psi_{l'} | U | \psi_k \rangle \langle \psi_k | \rho | \psi_l \rangle \text{Tr}(\Delta^{z_k - z_l} \sigma)$. The latter can be rewritten as $\text{Tr}(U^\dagger H_S U R_\sigma^{H_S}(\rho))$. \square

Given an Hermitian operator Q , we let $\lambda^\downarrow(Q)$ denote the collection of eigenvalues of Q (counted with multiplicities) ordered non-increasingly, while $\lambda^\uparrow(Q)$ denotes the non-decreasing ordering. The following lemma is a direct consequence of Theorem 4.3.53 on p. 255 in [51].

Lemma 16. *Let A, B be Hermitian operators on some finite-dimensional Hilbert space \mathcal{H} . Then*

$$\inf_{U \in \mathbb{U}(\mathcal{H})} \text{Tr}(U^\dagger A U B) = \sum_k \lambda_k^\downarrow(A) \lambda_k^\uparrow(B). \quad (\text{E4})$$

Lemma 17. *Let $s, \beta > 0$ and $M \in \mathbb{N}$ be given. Let \mathcal{H}_S be a finite-dimensional Hilbert space, and $H_S \in H_s(\mathcal{H}_S)$.*

Let $\rho \in \mathcal{S}(\mathcal{H}_S)$ and $\sigma \in \mathcal{S}^(\mathcal{H}_E)$. Then*

$$\begin{aligned} & \sup_{U \in \mathbb{U}(\mathcal{H}_S)} \text{Tr}([\hat{1}_S \otimes H_E^{(s)}]V(U)\rho \otimes \sigma V(U)^\dagger) \\ &= \text{Tr}(H_E^{(s)}\sigma) \\ & \quad + \frac{1}{\beta} D(\rho \| G(H_S)) \\ & \quad - \frac{1}{\beta} [S(R_\sigma^{H_S}(\rho)) - S(\rho)] \\ & \quad - \frac{1}{\beta} D\left(\lambda^\downarrow(R_\sigma^{H_S}(\rho)) \parallel \lambda^\downarrow(G(H_S))\right) \end{aligned} \quad (\text{E5})$$

Furthermore, there exists an element in $\mathbb{U}(\mathcal{H}_S)$ that achieves the above maximum, i.e., the ‘sup’ in the above equation can be replaced by ‘max’.

Compared to (D5) we recognize two new features in equation (E5). Firstly, the term $S(R_\sigma^{H_S}(\rho)) - S(\rho)$. Secondly, that we have $D(\lambda^\downarrow(R_\sigma^{H_S}(\rho)) \parallel \lambda^\downarrow(G(H_S)))$, rather than $D(\lambda^\downarrow(\rho) \parallel G(H_S))$. Both of these changes reflect the fact that the present model explicitly takes into account coherence. Due to Eq. (E2), in Lemma 14 it follows that $S(R_\sigma^{H_S}(\rho)) \geq S(\rho)$. Furthermore by the properties of relative entropy $D(\cdot \parallel \cdot) \geq 0$. Hence, the work yield is bounded from above by $kTD(\rho \parallel G(H_S))$. As we shall see in Section E 2, the term $S(R_\sigma^{H_S}(\rho)) - S(\rho)$ is determined by the relation between the degree of coherence in the energy reservoir, and the degree to which the state ρ contains superposition between energy eigenspaces. Analogously as to what we did in Section D 3, the term $D(\lambda^\downarrow(R_\sigma^{H_S}(\rho)) \parallel \lambda^\downarrow(G(H_S)))$ can be made small by introducing an ancillary Gibbs state.

Proof. By combining Lemmas 15 and 16 we obtain $\sup_U \text{Tr}([\hat{1}_S \otimes H_E^{(s)}]V(U)\rho \otimes \sigma V(U)^\dagger) = \text{Tr}(H_E^{(s)}\sigma) + \text{Tr}(H_S\rho) - \sum_k \lambda_k^\uparrow(H_S) \lambda_k^\downarrow(R_\sigma^{H_S}(\rho))$. Next we use $H_S = F(H_S) - kT \ln G(H_S)$, and $\lambda_k^\uparrow(H_S) = F(H_S) - kT \ln \lambda_k^\uparrow(G(H_S))$. The rest of the proof is merely a rearrangement to the terms.

Finally we should show that the ‘sup’ can be replaced by ‘max’. Let $G_n^\downarrow(H_S)$ be the eigenvalues of $G(H_S)$ ordered non-increasingly, and let $|\psi_n^\downarrow\rangle$ be corresponding eigenvectors. Let $|\phi_n^\downarrow\rangle$ be eigenvectors of $R_\sigma^{H_S}(\rho)$ corresponding to the eigenvalues of $\lambda^\downarrow(R_\sigma^{H_S}(\rho))$. The unitary operator $\tilde{U} := \sum_{n=1}^N |\psi_n^\downarrow\rangle \langle \phi_n^\downarrow|$ achieves the right hand side of equation (E5). \square

b. Adding a heat bath

Analogously to what we did in Section D 2, we here model the heat bath by appending an ancillary system in a Gibbs state.

As in Section D 2 we assume that the system S and the ancilla A are non-interacting, i.e., the total Hamiltonian is $H_{SA} = H_S \otimes \hat{1}_A + \hat{1}_S \otimes H_A$. (For the sake of notational simplicity we will in the following merely write $H_S + H_A$.)

We are allowed to freely choose the Hamiltonian H_A of the ancillary system, up to the condition that the eigenvalues have to be multiples of the energy-level spacing s in the energy reservoir, i.e., $H_A \in H_s(\mathcal{H}_A)$. We furthermore assume that the ancillary system always starts in the Gibbs state $G(H_A)$, and that the total initial state of the joint system is $\rho \otimes G(H_A)$.

Lemma 18.

Let $s > 0$ and let \mathcal{H}_S and \mathcal{H}_A be finite-dimensional Hilbert spaces and $H_S \in H_s(\mathcal{H}_S)$, $H_A \in H_s(\mathcal{H}_A)$. Let $\sigma \in \mathcal{S}^*(\mathcal{H}_E)$, then

$$R_\sigma^{H_S+H_A}(\rho \otimes G(H_A)) = R_\sigma^{H_S}(\rho) \otimes G(H_A), \quad (\text{E6})$$

for every $\rho \in \mathcal{S}(\mathcal{H}_S)$.

Proof. We let $h_n^S = sz_n^S$ and $|\psi_n^S\rangle$ be the eigenvalues and eigenvectors of H_S , and similarly let $h_m^A = sz_m^A$ and $|\psi_m^A\rangle$ be eigenvalues and eigenvectors of H_A . Hence, $h_n^S + h_m^A$ and $|\psi_n^S, \psi_m^A\rangle := |\psi_n^S\rangle|\psi_m^A\rangle$ are the eigenvalues and eigenvectors of $H_S + H_A$. The channel $R_\sigma^{H_S+H_A}$ can be written as

$$\begin{aligned} & R_\sigma^{H_S+H_A}(Q) \\ &= \sum_{n,n'=1}^N \sum_{m,m'=1}^N |\psi_n^S, \psi_m^A\rangle \langle \psi_n^S, \psi_m^A| Q |\psi_{n'}, \psi_{m'}^A\rangle \langle \psi_{n'}, \psi_{m'}^A| \\ & \quad \times \text{Tr}(\Delta^{z_n^S - z_{n'}^S + z_m^A - z_{m'}^A} \sigma). \end{aligned} \quad (\text{E7})$$

for all $Q \in L(\mathcal{H}_S \otimes \mathcal{H}_A)$. For $Q := \rho \otimes G(H_A)$ we obtain the right hand side of (E6) by using the fact that $G(H_A)$ is diagonal in $\{|\psi_m^A\rangle\}_m$. \square

By a direct combination of Lemma 17 and 18 we obtain the following.

Proposition 7. Let $s > 0$, \mathcal{H}_S and \mathcal{H}_A be finite-dimensional, and $H_S \in H_s(\mathcal{H}_S)$, $H_A \in H_s(\mathcal{H}_A)$. Let $\rho \in \mathcal{S}(\mathcal{H}_S)$ and $\sigma \in \mathcal{S}^*(\mathcal{H}_E)$. Then

$$\begin{aligned} & \sup_U \text{Tr}([\hat{1}_S \otimes \hat{1}_A \otimes H_E^{(s)}] V(U) \rho \otimes G(H_A) \otimes \sigma V(U)^\dagger) \\ &= \text{Tr}(H_E^{(s)} \sigma) \\ &+ \frac{1}{\beta} D(\rho \| G(H_S)) \\ &- \frac{1}{\beta} [S(R_\sigma^{H_S}(\rho)) - S(\rho)] \\ &- \frac{1}{\beta} D\left(\lambda^\downarrow(R_\sigma^{H_S}(\rho) \otimes G(H_A)) \middle\| \lambda^\downarrow(G(H_S) \otimes G(H_A))\right), \end{aligned} \quad (\text{E8})$$

where the supremum is taken over all elements $U \in U(\mathcal{H}_S)$.

Furthermore, there exists an element in $U(\mathcal{H}_S)$ that achieves the above equality, i.e., ‘sup’ can be replaced by ‘max’.

In comparison with (D6) we see that we again can affect the term $D(\lambda^\downarrow(\cdot) \| \lambda^\downarrow(\cdot))$ by adding an ancillary Gibbs state, while the other terms (including $S(R_\sigma^{H_S}(\rho)) - S(\rho)$) remain unaffected.

2. Coherence in expected work extraction

In light of Section A 3 it perhaps comes as no surprise that the level of coherence in the energy reservoir affects our ability to extract energy from the system. Here we first consider some natural special cases, and then proceed to bound $S(R_\sigma^{H_S}(\rho)) - S(\rho)$ in terms of the coherence in the energy reservoir.

a. Diagonal system states and diagonal energy-reservoir states

Let $H_S = \sum_l \tilde{h}_l P_l$ be the eigenvalue decomposition of H_S where P_l is the projector onto the eigenspace corresponding to eigenvalue \tilde{h}_l , where we let \tilde{h}_l denote the distinct eigenvalues of H_S , i.e., $\tilde{h}_l \neq \tilde{h}_{l'}$ if $l \neq l'$ (as opposed to h_n which is the complete list of eigenvalues including repetitions). We define the following operation

$$[Q]_H := \sum_l P_l Q P_l, \quad \forall Q \in L(\mathcal{H}). \quad (\text{E9})$$

In words $[Q]_H$ removes the off-diagonal blocks of Q with respect to the eigenspaces of H (i.e., $[\cdot]_H$ is what sometimes is referred to as a ‘pinching’). One can rewrite the channel $R_\sigma^{H_S}$ in terms of the eigenprojectors P_l as

$$R_\sigma^{H_S}(Q) = \sum_{l,l'} P_l Q P_{l'} \text{Tr}(\Delta^{(\tilde{h}_l - \tilde{h}_{l'})/s} \sigma). \quad (\text{E10})$$

By using this reformulation of $R_\sigma^{H_S}$ one can prove the following lemma.

Lemma 19. If $Q \in \mathcal{L}(\mathcal{H}_S)$ is such that $Q = [Q]_{H_S}$ then

$$R_\sigma^{H_S}(Q) = Q, \quad \forall \sigma \in \mathcal{S}(\mathcal{H}_E). \quad (\text{E11})$$

If σ is such that $\sigma = [\sigma]_{H_E^{(s)}}$, then

$$R_\sigma^{H_S}(Q) = [Q]_{H_S}, \quad \forall Q \in \mathcal{L}(\mathcal{H}_S). \quad (\text{E12})$$

By comparing Lemma 19 with Proposition 7 we can see that if ρ already is diagonal with respect to an energy eigenbasis, then $R_\sigma^{H_S}(\rho) = \rho$ and thus the term $S(R_\sigma^{H_S}(\rho)) - S(\rho)$ drops out. Furthermore, we obtain the following lemma, which tells us that if the state of the energy reservoir is diagonal, then the expected energy yield can not depend on the off-diagonal elements of the initial state ρ . This implies that if the reservoir is diagonal, the expected work extraction can only depend on $[\rho]_{H_S}$, which confirms the finding in [19].

Corollary 3. Let $s > 0$, \mathcal{H}_S and \mathcal{H}_A be finite-dimensional and $H_S \in H_s(\mathcal{H}_S)$, $H_A \in H_s(\mathcal{H}_A)$. Let $\rho \in \mathcal{S}(\mathcal{H}_S)$. If $\sigma \in \mathcal{S}^*(\mathcal{H}_E)$ is such that $\sigma = [\sigma]_{H_E^{(s)}}$, then

$$\begin{aligned} & \sup_U \text{Tr}([\hat{1}_S \otimes \hat{1}_A \otimes H_E^{(s)}]V(U)\rho \otimes G(H_A) \otimes \sigma V(U)^\dagger) \\ &= \text{Tr}(H_E^{(s)}\sigma) \\ &+ \frac{1}{\beta} D([\rho]_{H_S} \| G(H_S)) \\ &- \frac{1}{\beta} D\left(\lambda^\downarrow([\rho]_{H_S} \otimes G(H_A)) \left\| \lambda^\downarrow(G(H_S) \otimes G(H_A))\right.\right), \end{aligned} \quad (\text{E13})$$

where the supremum is taken over all elements $U \in U(H_S + H_A)$.

b. Bounds on $S(R_\sigma^{H_S}(\rho)) - S(\rho)$

As seen from the previous sections, the ‘quality’ of the work extraction is partially determined by how much the effective channel $R_\sigma^{H_S}$ increases the entropy of the initial state ρ . Here we determine bounds on this quantity in terms of the coherence in the energy reservoir.

Proposition 8. Let $\delta > 0$ and $M \in \mathbb{N}$. Let $\dim \mathcal{H}_S = N$, and let $H_S \in H_\delta(\mathcal{H}_S)$ with eigenvectors $|\psi_k\rangle$ and corresponding eigenvalues δz_k . Let $\sigma \in \mathcal{S}(\mathcal{H}_E)$. Let

$$\mathbf{F} := [1 - \text{Tr}(\Delta^{z_f - z_{k'}}\sigma)]_{k,k'=1}^N, \quad (\text{E14})$$

then

$$\sup_{\rho \in \mathcal{S}(\mathcal{H}_S)} \|\rho - R_\sigma^{H_S}(\rho)\|_1 \leq \|\mathbf{F}\|. \quad (\text{E15})$$

Proof. By using (E1) in Definition 2

$$\rho - R_\sigma^{H_S}(\rho) = \sum_{k,k'=1}^N |\psi_k\rangle\langle\psi_k|\rho|\psi_{k'}\rangle\langle\psi_{k'}|\mathbf{F}_{k,k'}. \quad (\text{E16})$$

If we define the matrix $\boldsymbol{\rho} = [|\psi_k\rangle\langle\psi_k|\rho|\psi_{k'}\rangle\langle\psi_{k'}|]_{k,k'=1}^N$ it follows by equation (A17) that $\|\rho - R_\sigma^{H_S}(\rho)\|_1 = \|\boldsymbol{\rho} * \mathbf{F}\|_1 \leq \|\boldsymbol{\rho}\|_1 \|\mathbf{F}\| = \|\mathbf{F}\|$. \square

Lemma 20. Let \mathcal{H}_S be finite-dimensional with $N = \dim(\mathcal{H}_S) \geq 2$, and let $H_S \in H_s(\mathcal{H}_S)$, with eigenvalues $s z_m$. Let $|\eta_{L,l_0}\rangle := \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} |l + l_0\rangle \in \mathcal{H}_E$, then

$$\sup_{\rho \in \mathcal{S}(\mathcal{H}_S)} \|\rho - R_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|}^{H_S}(\rho)\|_1 \leq \frac{N(z_{\max} - z_{\min})}{L}, \quad (\text{E17})$$

where $z_{\max} := \max_{k=1}^N z_k$ and $z_{\min} := \min_{k=1}^N z_k$.

Proof. In the case $L < z_{\max} - z_{\min}$, Eq. (E17) is trivially true. This follows since both ρ and $R_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|}^{H_S}(\rho)$ are density operators and thus $\|\rho - R_{|\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|}^{H_S}(\rho)\| \leq 2 \leq$

N . We can thus without loss of generality assume that $L \geq z_{\max} - z_{\min}$.

Due to the latter inequality we can conclude [e.g., via (A14)] that $\mathbf{F}_{k,k'} = |z_k - z_{k'}|/L$. By combining this observation with Proposition 8 and Lemma 4 we can conclude that $\|\rho - R_\sigma^{H_S}(\rho)\|_1 \leq \|\mathbf{F}\| \leq N(z_{\max} - z_{\min})/L$. \square

The following bound for the von Neumann entropy is proved in [52]

Lemma 21 ([52]). Let $\dim \mathcal{H} = N$ and $\rho, \sigma \in \mathcal{S}(\mathcal{H})$, then

$$\begin{aligned} |S(\rho) - S(\sigma)| &\leq \frac{1}{2} \|\rho - \sigma\|_1 \log(N-1) + \Xi\left(\frac{1}{2} \|\rho - \sigma\|_1\right), \\ \Xi(x) &:= -x \ln x - (1-x) \ln(1-x). \end{aligned} \quad (\text{E18})$$

Taking into account the fact that Ξ is monotonically increasing on the interval $[0, 1/2]$ and thus $\Xi(\|\rho - \sigma\|_1/2) \leq \Xi(N(z_{\max} - z_{\min})/(2L))$ if $\|\rho - \sigma\|_1 \leq N(z_{\max} - z_{\min})/L \leq 1$ we obtain the following bound by combining Lemmas 20 and 21.

Proposition 9. Let $s > 0$ and let \mathcal{H}_S be a Hilbert space with $\dim \mathcal{H}_S = N$, and let $H_S \in H_s(\mathcal{H}_S)$. Let z_{\max} be the largest eigenvalue of H_S divided by s , and let z_{\min} be the smallest eigenvalue of H_S divided by s . Assume $L \geq N(z_{\max} - z_{\min})$, and define

$$\sigma_L := |\eta_{L,l_0}\rangle\langle\eta_{L,l_0}|, \quad |\eta_{L,l_0}\rangle := \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} |l + l_0\rangle_E. \quad (\text{E19})$$

Then

$$\begin{aligned} 0 \leq S(R_{\sigma_L}^{H_S}(\rho)) - S(\rho) &\leq \frac{z_{\max} - z_{\min}}{2L} N \log N \\ &+ \Xi\left(\frac{N(z_{\max} - z_{\min})}{2L}\right), \end{aligned} \quad (\text{E20})$$

with Ξ as in Lemma 21.

3. Achieving the ‘standard’ optimal expected work extraction

Here we show that we can regain the ‘standard’ result on expected work extraction, namely that the maximum yield is $kTD(\rho \| G(H))$. We will obtain this as a limit case, namely of a large density of states in the energy reservoir, high degree of coherence in the energy reservoir, and very large (in Hilbert space dimension) ancillary systems.

To do this, we introduce a number J that controls the density of states in the reservoir, compared to the energy levels in S . More precisely, if the system Hamiltonian H_S has eigenvalues that are multiples of δ , then we assume that the energy spacings in the energy ladder of the reservoir is $s := \delta/J$. To provide coherence, we assume our standard uniform superpositions

$|\eta_{L,l_0}\rangle = \sum_{l=0}^{L-1} |l + l_0\rangle / \sqrt{L}$, and thus get a number L that in some sense controls the degree of coherence in the reservoir.

To construct the above mentioned limit, we consider a fixed Hamiltonian $H_S \in \mathcal{H}_\delta(\mathcal{H}_S)$, but a sequence of energy reservoirs, one for each $K \in \mathbb{N}$ (where K is just a ‘dummy index’). We assume that the K th reservoir has the energy spacing $s(K) := \delta/J(K)$ and the initial state $|\eta_{L(K),l_0}\rangle$ (where l_0 does not matter), with $J(K)$ and $L(K)$ being functions of K . We furthermore construct a sequence of ancillary systems A_K with Hilbert spaces \mathcal{H}_A^K with dimension N^{K+1} , where $N = \dim \mathcal{H}_S$. These ancillary systems are Gibbs distributed with respect to suitably chosen Hamiltonians.

It maybe should be noted that in these derivations we limit ourselves to states ρ that have full support ($\rho \in \mathcal{S}_+(\mathcal{H}_S)$). This is essentially due to somewhat too weak bounds, which involves the logarithm of the smallest eigenvalue of ρ . Reasonably it should be possible to overcome this, but we will not do this here.

Corollary 4. *Let $N \in \mathbb{N}$, $\delta, \beta > 0$ be given. Let \mathcal{H}_S be a Hilbert space with $\dim \mathcal{H}_S = N$, let $\rho \in \mathcal{S}_+(\mathcal{H}_S)$, and $H_S \in \mathcal{H}_\delta(\mathcal{H}_S)$. Let $L, J : \mathbb{N} \rightarrow \mathbb{N}$ be such that $\lim_{K \rightarrow +\infty} [K/J(K)] = 0$ and $\lim_{K \rightarrow +\infty} [J(K)/L(K)] = 0$. Then there exist a sequence of ancillary Hilbert spaces \mathcal{H}_A^K with $\dim \mathcal{H}_A^K = N^{K+1}$, and $H_A^K \in \mathcal{H}_{\delta/J(K)}(\mathcal{H}_A^K)$ such that*

$$\lim_{K \rightarrow \infty} W_{K,J(K),L(K)} = \frac{1}{\beta} D(\rho \| G(H_S)), \quad (\text{E21})$$

where

$$W_{K,J(K),L(K)} := \sup_{U \in U(\mathcal{H}_S \otimes \mathcal{H}_A^K)} \text{Tr}(H_E^{\delta/J(K)} \tilde{\sigma}) - \text{Tr}(H_E^{\delta/J(K)} \sigma_{L(K)}) \quad (\text{E22})$$

$$\tilde{\sigma} := \text{Tr}_{S,A}[V(U)\rho \otimes G(H_A^K) \otimes \sigma_L V(U)^\dagger] \quad (\text{E23})$$

and where $V(U)$ is as defined in (A5) for the reservoir Hamiltonian $H_E^{(\delta/J(K))}$, and

$$\begin{aligned} \sigma_{L(K)} &:= |\eta_{L(K),l_0}\rangle \langle \eta_{L(K),l_0}|, \\ |\eta_{L(K),l_0}\rangle &:= \frac{1}{\sqrt{L(K)}} \sum_{l=0}^{L(K)-1} |l + l_0\rangle_E, \end{aligned} \quad (\text{E24})$$

for some fixed $l_0 \in \mathbb{Z}$.

This is a corollary of Proposition 11. The following subsections are devoted to the proof of this Proposition.

It is maybe worth to point out that although the specific construction we use to prove this proposition does require both increasing sizes of the ancillary systems, as well as increasing density of states in the energy reservoir, we do not strictly speaking prove that these are necessary conditions (although the author suspects them to be necessary in the generic case). Note also that this increase in density of states is in line with what was found in [19]. Recent results [53] may help to determine whether these are also necessary conditions.

a. Shifting the path to the grid

In Appendix D3 we demonstrated that one can obtain a discretized path of probability distributions, such that the sum of all consecutive relative entropies along the path can be made arbitrarily small. These probability distributions we constructed as Gibbs distributions $G(h^k)$ of underlying energy level configurations h^k . The problem is that these energy level configurations may not necessarily satisfy our energy matching criterion, i.e., that each h_n^k should be a multiple of the energy reservoir spacing s . The purpose of this Section is to show that we can shift the energy level configurations h^k into new configurations \tilde{h}^k that are multiples of s , but at the same time obtain a sum of consecutive relative entropies $\sum_{k=0}^{K-1} D(G(\tilde{h}^k) \| G(\tilde{h}^{k+1}))$ that is sufficiently close to the original $\sum_{k=0}^{K-1} D(G(h^k) \| G(h^{k+1}))$. In \mathbb{R}^N , regarded as the space of energy level configurations h , the configurations that respect the energy matching form a square lattice. We construct the new configuration \tilde{h} by $\tilde{h}_n = s \lfloor h_n/s \rfloor$, where $\lfloor \cdot \rfloor$ is the floor-function, which rounds of real numbers to the nearest smaller integer. (There is no particular reason to choose the floor-function in particular. We could equally well use $\lceil \cdot \rceil$, or some other method to shift the original point h to somewhere close in the lattice.)

Lemma 22. *Let $N, K \in \mathbb{N}$ and $\beta, s > 0$ be given. Let $(h^k)_{k=0}^K$ with $h^k \in \mathbb{R}^N$. Then there exists $(\tilde{h}^k)_{k=0}^K$ with $\tilde{h}^k \in s\mathbb{Z}^N$, such that*

$$\begin{aligned} \max_{n=1,\dots,N} |\tilde{h}_n^k - h_n^k| &\leq s, \quad D(G(h^k) \| G(\tilde{h}^k)) \leq s\beta, \\ D(G(\tilde{h}^k) \| G(h^k)) &\leq s\beta, \end{aligned} \quad (\text{E25})$$

for $k = 0, \dots, K$, and such that

$$\begin{aligned} \left| \sum_{k=0}^{K-1} D(G(\tilde{h}^k) \| G(\tilde{h}^{k+1})) - \sum_{k=0}^{K-1} D(G(h^k) \| G(h^{k+1})) \right| \\ \leq 2\beta s(K+1) \\ + \beta(e^{s\beta} - 1)K \max_{k=0,\dots,K-1} \max_{n=1,\dots,N} |h_n^{k+1} - h_n^k|. \end{aligned} \quad (\text{E26})$$

Proof. For each k let $\tilde{h}_n^k := s \lfloor \frac{1}{s} h_n^k \rfloor$, for $n = 1, \dots, N$. By construction, this means that $\tilde{h}^k \in s\mathbb{Z}^N$, and

$$\begin{aligned} |\tilde{h}_n^k - h_n^k| &\leq s, \quad \beta s > \ln \frac{Z(\tilde{h}^k)}{Z(h^k)} \geq 0, \\ e^{s\beta} &> \frac{G_n(\tilde{h})}{G_n(h)} \geq e^{-s\beta}. \end{aligned} \quad (\text{E27})$$

One can moreover realize that

$$|G_n(\tilde{h}) - G_n(h)| \leq G_n(h)(e^{s\beta} - 1). \quad (\text{E28})$$

By Eq. (E27) it follows that $s\beta \geq D(G(h) \| G(\tilde{h}))$. By the analogous reasoning one finds $s\beta \geq D(G(\tilde{h}) \| G(h))$.

Observe that

$$\begin{aligned} & \sum_{k=0}^{K-1} D(G(h^k) \| G(h^{k+1})) \\ &= \beta \sum_{k=0}^{K-1} \sum_{n=1}^N G_n(h^k) (h_n^{k+1} - h_n^k) + \ln Z(h^K) - \ln Z(h^0) \end{aligned} \quad (\text{E29})$$

and the analogous statement for \tilde{h} . By making use of the inequalities in (E27) we find

$$\begin{aligned} & \left| \sum_{k=0}^{K-1} D(G(\tilde{h}^k) \| G(\tilde{h}^{k+1})) - \sum_{k=0}^{K-1} D(G(h^k) \| G(h^{k+1})) \right| \\ & \leq \left| \beta \sum_{k=0}^{K-1} \sum_{n=1}^N G_n(\tilde{h}^k) (\tilde{h}_n^{k+1} - \tilde{h}_n^k) \right. \\ & \quad \left. - \beta \sum_{k=0}^{K-1} \sum_{n=1}^N G_n(h^k) (h_n^{k+1} - h_n^k) \right| + 2\beta s \end{aligned} \quad (\text{E30})$$

$$\begin{aligned} &= \left| \beta \sum_{k=0}^{K-1} \sum_{n=1}^N G_n(\tilde{h}^k) (\tilde{h}_n^{k+1} - h_n^{k+1} - \tilde{h}_n^k + h_n^k) \right. \\ & \quad \left. + \beta \sum_{k=0}^{K-1} \sum_{n=1}^N [G_n(\tilde{h}^k) - G_n(h^k)] (h_n^{k+1} - h_n^k) \right| \\ & \quad + 2\beta s \\ & \leq 2\beta s(K+1) \\ & \quad + \beta \sum_{k=0}^{K-1} \sum_{n=1}^N |G_n(\tilde{h}^k) - G_n(h^k)| |h_n^{k+1} - h_n^k| \\ & \quad [\text{By Eq. (E28)}] \end{aligned}$$

$$\begin{aligned} & \leq 2\beta s(K+1) + \beta(e^{s\beta} - 1) \sum_{k=0}^{K-1} \sum_{n=1}^N G_n(h^k) |h_n^{k+1} - h_n^k| \\ & \leq 2\beta s(K+1) + \beta(e^{s\beta} - 1) K \max_{k=0}^{K-1} \max_{n=1, \dots, N} |h_n^{k+1} - h_n^k| \end{aligned} \quad (\text{E31})$$

□

b. Simulation of the path

Here we remind that $\mathbb{P}(N)$ denotes the set of probability distributions over N symbols, and $\mathbb{P}_+(N)$ denotes the subset of probability distributions with full support.

Proposition 10. *Let $N \in \mathbb{N}$, $\delta, \beta > 0$, and $q, p \in \mathbb{P}_+(N)$ be given. For each $K, J \in \mathbb{N}$ there exists a $h^K \in \frac{\delta}{J} \mathbb{Z}^{N^{K+1}}$,*

such that

$$\begin{aligned} & D\left((qG(h^K))^\downarrow \| (pG(h^K))^\downarrow\right) \\ & \leq 4\beta\delta \frac{1}{J} + 2\beta\delta \frac{K}{J} \\ & \quad + (e^{\frac{\delta\beta}{J}} - 1) \max_{n=1, \dots, N} \left| \ln \frac{q_n}{p_n} \right| \\ & \quad + \frac{1}{K} \max_n \left| \ln \frac{q_n}{p_n} \right|^2. \end{aligned} \quad (\text{E32})$$

Proof. Since $q, p \in \mathbb{P}_+(N)$, it follows that $\log q_n$ and $\log p_n$ are well defined and finite. Thus we can define $h^i := -\frac{1}{\beta} \ln q$, $h^f := -\frac{1}{\beta} \ln p$. We furthermore define $h : [0, 1] \rightarrow \mathbb{R}^N$ by $h(x) := (1-x)h^i + xh^f$ for $x \in [0, 1]$. The function $h(x) = (1-x)h^i + xh^f$ clearly has a continuous second derivative on each component.

For this function, let $h^k := h(\frac{k}{K})$, $s := \frac{\delta}{J}$ in Lemma 22. Lemma 22 thus delivers $\tilde{h}^{k:K} := \tilde{h}^k$, with $\tilde{h}^{k:K} \in \frac{\delta}{J} \mathbb{Z}^N$, such that

$$\begin{aligned} 0 & \leq D(G(h(0)) \| G(h^{0:K})) \leq \frac{\delta\beta}{J}, \\ 0 & \leq D(G(\tilde{h}^{K:K}) \| G(h(1))) \leq \frac{\delta\beta}{J}, \end{aligned} \quad (\text{E33})$$

and

$$\begin{aligned} & \left| \sum_{k=0}^{K-1} D(G(\tilde{h}^{k:K}) \| G(\tilde{h}^{k+1:K})) \right. \\ & \quad \left. - \sum_{k=0}^{K-1} D(G(h(\frac{k}{K})) \| G(h(\frac{k+1}{K}))) \right| \\ & \leq 2\beta\delta \frac{K+1}{J} + (e^{\frac{\delta\beta}{J}} - 1) \max_{n=1, \dots, N} \left| \ln \frac{q_n}{p_n} \right| \end{aligned} \quad (\text{E34})$$

A combination of (E34) and Lemma 11 yields

$$\begin{aligned} 0 & \leq \sum_{k=0}^{K-1} D(G(\tilde{h}^{k:K}) \| G(\tilde{h}^{k+1:K})) \\ & \leq \left| \sum_{k=0}^{K-1} D(G(\tilde{h}^{k:K}) \| G(\tilde{h}^{k+1:K})) \right. \\ & \quad \left. - \sum_{k=0}^{K-1} D(G(h(\frac{k}{K})) \| G(h(\frac{k+1}{K}))) \right| \\ & \quad + \sum_{k=0}^{K-1} D(G(h(\frac{k}{K})) \| G(h(\frac{k+1}{K}))) \\ & \leq 2\beta\delta \frac{K+1}{J} + (e^{\frac{\delta\beta}{J}} - 1) \max_{n=1, \dots, N} \left| \ln \frac{q_n}{p_n} \right| \\ & \quad + \frac{1}{K} \max_{n=1, \dots, N} \left| \ln \frac{q_n}{p_n} \right|^2 \end{aligned} \quad (\text{E35})$$

By combining this with the two inequalities in Eq. (E33)

we thus obtain

$$\begin{aligned}
0 &\leq D\left(G(h(0))\right)\left\|G(\tilde{h}^{0:K})\right\| \\
&\quad + \sum_{k=0}^{K-1} D\left(G(\tilde{h}^{k:K})\right)\left\|G(\tilde{h}^{k+1:K})\right\| \\
&\quad + D\left(G(\tilde{h}^{K:K})\right)\left\|G(h(1))\right\| \\
&\leq 4\beta\delta\frac{1}{J} + 2\beta\delta\frac{K}{J} \\
&\quad + (e^{\frac{\delta\beta}{J}} - 1) \max_{n=1,\dots,N} \left|\ln \frac{q_n}{p_n}\right| \\
&\quad + \frac{1}{K} \max_{n=1,\dots,N} \left|\ln \frac{q_n}{p_n}\right|^2.
\end{aligned} \tag{E36}$$

Define

$$\begin{aligned}
h_{n_0,\dots,n_K}^K &:= \tilde{h}_{n_0}^{0:K} + \dots + \tilde{h}_{n_K}^{K:K}, \\
(n_0, \dots, n_K) &\in \{1, \dots, N\}^{K+1}.
\end{aligned} \tag{E37}$$

Since $h^{k:K} \in \frac{\delta}{J}\mathbb{Z}^N$ it follows that

$$h^K \in \frac{\delta}{J}\mathbb{Z}^{N^{K+1}}. \tag{E38}$$

Furthermore

$$G_{n_0,\dots,n_K}(h^K) = G_{n_0}(\tilde{h}^{0:K}) \dots G_{n_{K-1}}(\tilde{h}^{K-1:K}) G_{n_K}(\tilde{h}^{K:K}). \tag{E39}$$

On $\mathbb{R}^{N^{K+2}}$ we define the operation Π by

$$(\Pi a)_{n_s n_0, n_1, \dots, n_{K-1}, n_K} := a_{n_0, n_1, \dots, n_{K-1}, n_K, n_s}. \tag{E40}$$

In other words, Π is defined via a cyclic permutation of the index of the indices of a . By Corollary 2 it follows that

$$\begin{aligned}
&D\left((qG(h^K))^\downarrow\right)\left\|pG(h^K)^\downarrow\right\| \\
&\leq D\left(\Pi(qG(h^K))\right)\left\|pG(h^K)\right\| \\
&= D\left(\Pi(qG(\tilde{h}^{0:K}) \dots G(\tilde{h}^{K-1:K}) G(\tilde{h}^{K:K}))\right)\left\|pG(\tilde{h}^{0:K}) \dots G(\tilde{h}^{K-1:K}) G(\tilde{h}^{K:K})\right\| \\
&= D\left(G(\tilde{h}^{K:K}) qG(\tilde{h}^{0:K}) \dots G(\tilde{h}^{K-1:K})\right)\left\|pG(\tilde{h}^{0:K}) \dots G(\tilde{h}^{K-1:K}) G(\tilde{h}^{K:K})\right\| \\
&= D(q\|G(\tilde{h}^{0:K})) \\
&\quad + \sum_{k=0}^{K-1} D(G(\tilde{h}^{k:K})\|G(\tilde{h}^{k+1:K})) \\
&\quad + D(G(\tilde{h}^{K:K})\|p).
\end{aligned} \tag{E41}$$

By inserting (E35) we obtain the statement of the proposition. \square

c. Assembling it all

A channel $\Gamma : L(\mathcal{H}) \rightarrow L(\mathcal{H})$ is unital (or mixing enhancing) if $\Gamma(\hat{1}) = \hat{1}$. For a positive operator Q , let $\lambda_{\min}(Q)$ denote the minimal eigenvalue.

Lemma 23. *If a channel Γ is unital then $\lambda_{\min}(\Gamma(\rho)) \geq \lambda_{\min}(\rho)$ for all $\rho \in \mathcal{S}(\mathcal{H})$.*

Proof. By [54], or Theorem 2 in [55], we know that for unital channels the output is more mixed than the input, i.e., $\lambda(\Gamma(\rho)) \prec \lambda(\rho)$. If the underlying Hilbert space has dimension N , this means $\sum_{k=1}^n \lambda_k^\downarrow(\Gamma(\rho)) \leq \sum_{k=1}^n \lambda_k^\downarrow(\rho)$ for $n = 1, \dots, N$. Since Γ is a channel and thus trace preserving, we can conclude that $\lambda_{\min}(\Gamma(\rho)) = 1 - \sum_{k=1}^{n-1} \lambda_k^\downarrow(\Gamma(\rho)) \geq 1 - \sum_{k=1}^{n-1} \lambda_k^\downarrow(\rho) = \lambda_{\min}(\rho)$. \square

To avoid confusion in relation to Corollary 4, one should note that in both Proposition 11 and 12, J and L are not regarded as functions of K .

Note furthermore the change of notation from $R_\sigma^{H_S}$ to $R_\sigma^{\delta/J, H_S}$. This serves to underline the fact that $R_\sigma^{\delta/J, H_S}$ depends on J , which should be kept in mind when using Proposition 11. For a fixed density operator $\sigma = \sum_{jj'} \sigma_{jj'} |j\rangle\langle j'|$, an increase of J effectively means a ‘decrease’ of the coherence in σ compared to H_S . To see this, consider a transition from $|\psi_n\rangle$ to $|\psi_{n'}\rangle$. The corresponding change in energy, $s(z_{n'} - z_n)$, has to be compensated for by a change of energy $s(z_n - z_{n'}) = sJ(x_n - x_{n'})$ in the reservoir. The larger the J , the larger the compensating jumps in the reservoir have to be (counted in numbers of eigenstates). In other words, an increase in J effectively decreases the ‘width’ of σ in relation to H_S , and in this sense means a decrease in coherence.

Proposition 11. *Let $N \in \mathbb{N}$, $\delta, \beta > 0$ be given. Let \mathcal{H}_S be a Hilbert space with $\dim \mathcal{H}_S = N$, let $\rho \in \mathcal{S}_+(\mathcal{H}_S)$, and $H_S \in H_\delta(\mathcal{H}_S)$, and let $\sigma \in \mathcal{S}^*(\mathcal{H}_E)$. Then, for every $J, K \in \mathbb{N}$ there exists an ancillary Hilbert space \mathcal{H}_A^K with $\dim \mathcal{H}_A^K = N^{K+1}$, and $H_A^K \in H_{\delta/J}(\mathcal{H}_A^K)$ such that*

$$\begin{aligned}
0 &\leq \frac{1}{\beta} D(\rho\|G(H_S)) - \frac{1}{\beta} \left[S(R_\sigma^{\delta/J, H_S}(\rho)) - S(\rho) \right] - W_{J,K}(\sigma) \\
&\leq 4\delta\frac{1}{J} + 2\delta\frac{K}{J} \\
&\quad + \frac{1}{\beta} (e^{\frac{\delta\beta}{J}} - 1) |\ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))]| \\
&\quad + \frac{1}{K\beta} \left| \ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))] \right|^2,
\end{aligned} \tag{E42}$$

where

$$\begin{aligned}
W_{J,K}(\sigma) &:= \sup_{U \in U(\mathcal{H}_S \otimes \mathcal{H}_A^K)} \text{Tr}(H_E^{(\delta/J)} \tilde{\sigma}) \\
&\quad - \text{Tr}(H_E^{(\delta/J)} \sigma), \\
\tilde{\sigma} &:= \text{Tr}_{S,A}[V(U)\rho \otimes G(H_A^K) \otimes \sigma V(U)^\dagger],
\end{aligned} \tag{E43}$$

and where $V(U)$ is as defined in (A5) and $R_\sigma^{\delta/J, H_S}$ is as in Definition 2, for the reservoir Hamiltonian $H_E^{(\delta/J)}$.

Proof. Since $R_\sigma^{\delta/J, H_S}$ is unital (Lemma 14) we can, by Lemma 23, conclude that

$$\lambda_{\min}(R_\sigma^{\delta/J, H_S}(\rho)) \geq \lambda_{\min}(\rho). \quad (\text{E44})$$

Hence, since ρ by assumption is full rank, it follows that $R_\sigma^{\delta/J, H_S}(\rho)$ is full rank. Let q denote the eigenvalues of $R_\sigma^{\delta/J, H_S}(\rho)$, i.e., $q := \lambda(R_\sigma^{\delta/J, H_S}(\rho))$. Let p denote the eigenvalues of $G(H_S)$, i.e., $p := \lambda(G(H_S))$. Since ρ by assumption is full rank, it follows that $R_\sigma^{\delta/J, H_S}(\rho)$ is full rank, and thus $q \in \mathbb{P}_+(N)$. Furthermore $p \in \mathbb{P}_+(N)$ since $G(H_S)$ is a Gibbs distribution. Furthermore,

$$\begin{aligned} \left| \ln \frac{q_n}{p_n} \right| &\leq -\ln \lambda_{\min}(R_\sigma^{\delta/J, H_S}(\rho)) - \ln \lambda_{\min}(G(H_S)) \\ &\leq -\ln \lambda_{\min}(\rho) - \ln \lambda_{\min}(G(H_S)), \end{aligned} \quad (\text{E45})$$

where the last inequality follows by (E44).

Since $q, p \in \mathbb{P}_+(N)$, we can conclude that the conditions of Proposition 10 are satisfied, and we thus know that for each K, J there exists a $h^K \in \frac{\delta}{J} \mathbb{Z}^{N^{K+1}}$ such that

$$\begin{aligned} &D\left((qG(h^K))^\downarrow \parallel (pG(h^K))^\downarrow\right) \\ &\leq 4\beta\delta \frac{1}{J} + 2\beta\delta \frac{K}{J} \\ &\quad + (e^{\frac{\delta\beta}{J}} - 1) |\ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))]| \\ &\quad + \frac{1}{K} \max_n |\ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))]|^2, \end{aligned} \quad (\text{E46})$$

where we have used (E45). Let \mathcal{H}_A^K be a Hilbert space with $\dim \mathcal{H}_A^K = N^{K+1}$. We can construct a Hermitian operator H_A^K on \mathcal{H}_A^K that has h^K as its eigenvalues. By construction it follows that $H_A^K \in H_{\delta/J}(\mathcal{H}_A^K)$.

One can furthermore check that

$$\begin{aligned} &D\left(\lambda^\downarrow(R_\sigma^{\delta/J, H_S}(\rho) \otimes G(H_A^K)) \parallel \lambda^\downarrow(G(H_S) \otimes G(H_A^K))\right) \\ &= D\left((qG(h^K))^\downarrow \parallel (pG(h^K))^\downarrow\right). \end{aligned} \quad (\text{E47})$$

By assumption $H_S \in H_\delta(\mathcal{H}_S)$, while $H_A^K \in H_{\delta/J}(\mathcal{H}_A^K)$ and the energy reservoir has the energy spacing $s = \delta/J$. Note that

$$H_\delta(\mathcal{H}_S) \subset H_{\delta/J}(\mathcal{H}_S). \quad (\text{E48})$$

The assumption $H_S \in H_\delta(\mathcal{H}_S)$ implies $H_S \in H_{\delta/J}(\mathcal{H}_S)$ and thus allows us to apply Proposition 7 on the state $\rho \otimes G(H_A^K) \otimes \sigma$, which thus yields

$$\begin{aligned} &W_{J,K}(\sigma) \\ &= \frac{1}{\beta} D(\rho \parallel G(H_S)) - \frac{1}{\beta} \left[S(R_\sigma^{\delta/J, H_S}(\rho)) - S(\rho) \right] \\ &\quad - \frac{1}{\beta} D\left(\lambda^\downarrow(R_\sigma^{\delta/J, H_S}(\rho) \otimes G(H_A^K)) \parallel \lambda^\downarrow(G(H_S) \otimes G(H_A^K))\right). \end{aligned} \quad (\text{E49})$$

By combining (E46), (E47), and (E49) we obtain the proposition. \square

As was mentioned above, an increase of J effectively means a decrease of the coherence in the reservoir state σ . To guarantee that $S(R_\sigma^{\delta/J, H_S}(\rho)) - S(\rho)$ is small, we must thus choose the state σ as to compensate for a large J . In the case $\sigma = |\eta_{L, l_0}\rangle\langle\eta_{L, l_0}|$, we must thus make sure that $L \gg J$.

Proposition 12. *Let $N \in \mathbb{N}$, $\delta, \beta > 0$ be given. Let \mathcal{H}_S be a Hilbert space with $\dim \mathcal{H}_S = N$, let $\rho \in \mathcal{S}_+(\mathcal{H}_S)$, and $H_S \in H_\delta(\mathcal{H}_S)$. With h_n^S the eigenvalues of H_S , we let $x_n := h_n/\delta$, and $x_{\min} := \min_n x_n$ and $x_{\max} := \max_n x_n$. Let $J, K, L \in \mathbb{N}$ be such that $L \geq NJ(x_{\max} - x_{\min})$, then there exists an ancillary Hilbert space \mathcal{H}_A^K with $\dim \mathcal{H}_A^K = N^{K+1}$, and $H_A^K \in H_{\delta/J}(\mathcal{H}_A^K)$ such that*

$$\begin{aligned} 0 &\leq \frac{1}{\beta} D(\rho \parallel G(H_S)) - W_{K,J,L} \\ &\leq \frac{J}{\beta L} \frac{x_{\max} - x_{\min}}{2} N \log N + \frac{1}{\beta} \Xi \left(\frac{J}{L} \frac{N(x_{\max} - x_{\min})}{2} \right) \\ &\quad + 4\delta \frac{1}{J} + 2\delta \frac{K}{J} \\ &\quad + \frac{1}{\beta} (e^{\frac{\delta\beta}{J}} - 1) |\ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))]| \\ &\quad + \frac{1}{K\beta} |\ln[\lambda_{\min}(\rho)\lambda_{\min}(G(H_S))]|^2, \end{aligned} \quad (\text{E50})$$

where

$$\begin{aligned} W_{K,J,L} &:= \sup_{U \in U(\mathcal{H}_S \otimes \mathcal{H}_A^K)} \text{Tr}(H_E^{\delta/J} \tilde{\sigma}) \\ &\quad - \text{Tr}(H_E^{\delta/J} \sigma_L) \end{aligned} \quad (\text{E51})$$

$$\tilde{\sigma} := \text{Tr}_{S,A}[V(U)\rho \otimes G(H_A^K) \otimes \sigma_L V(U)^\dagger] \quad (\text{E52})$$

and where $V(U)$ is as defined in (A5) for the reservoir Hamiltonian $H_E^{(\delta/J)}$, and

$$\begin{aligned} \sigma_L &:= |\eta_{L, l_0}\rangle\langle\eta_{L, l_0}|, \\ |\eta_{L, l_0}\rangle &:= \frac{1}{\sqrt{L}} \sum_{l=0}^{L-1} |l + l_0\rangle_E, \end{aligned} \quad (\text{E53})$$

for some fixed $l_0 \in \mathbb{Z}$.

By letting J and L be functions of the parameter K , we can use Proposition 12 to prove Corollary 4. For this one can use the fact that due to $\lim_{K \rightarrow +\infty} [J(K)/L(K)] = 0$, the inequality $L(K) \geq NJ(K)(x_{\max} - x_{\min})$ will hold for all sufficiently large K .

Proof of Proposition 12. The eigenvalues h_n^S of H_S can be expressed in terms of multiples of the energy spacing $s = \delta/J$ of the energy reservoir, as $h_n^S = z_n \delta/J$. It proves convenient to re-express h_n^S in terms of multiples in δ

(which we can do, due to the assumption $H_S \in H_\delta(\mathcal{H}_S)$). Hence, $h_n^S = x_n \delta$ and $z_n = x_n J$, and thus $z_{\max} = x_{\max} J$ and $z_{\min} = x_{\min} J$.

By assumption $L \geq NJ(x_{\max} - x_{\min})$, and thus $L \geq N(z_{\max} - z_{\min})$. Hence, Proposition 9 is applicable and yields

$$\begin{aligned} 0 &\leq S(R_{\sigma_L}^{\delta/J, H_S}(\rho)) - S(\rho) \\ &\leq \frac{z_{\max} - z_{\min}}{2L} N \log N + \Xi\left(\frac{N(z_{\max} - z_{\min})}{2L}\right) \\ &\leq \frac{J}{L} \frac{x_{\max} - x_{\min}}{2} N \log N + \Xi\left(\frac{JN}{L} \frac{x_{\max} - x_{\min}}{2}\right). \end{aligned} \quad (\text{E54})$$

If we let $\sigma := \sigma_L$ in Proposition 11 and define $W_{J,K,L} := W_{J,K}(\sigma_L)$, and combine this with (E54), we thus obtain (E50). \square

-
- [1] E. T. Jaynes, F. W. Cummings, Comparisons of quantum and semiclassical radiation theories with application to the beam maser.
 - [2] B. W. Shore, P. L. Knight, The Jaynes-Cummings model, *J. Mod. Opt.* **40**, 1195 (1993).
 - [3] L. Szilard, Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen, *Z. Phys.* **53**, 840 (1929).
 - [4] R. Landauer, Irreversibility and heat generation in the computing process, *IBM J. Res. Develop.* **5**, 183 (1961).
 - [5] I. Procaccia and R. D. Levine, Potential work: A statistical-mechanical approach for systems in disequilibrium, *J. Chem. Phys.* **65**, 3357 (1967).
 - [6] C. H. Bennett, Notes on Landauer's principle, reversible computation, and Maxwell's Demon, *Stud. Hist. Phil. Mod. Phys.* **34**, 501 (2003).
 - [7] H. S. Leff and A. F. Rex, *Maxwell's Demon: Entropy, Information, Computing* (Taylor and Francis, 1990).
 - [8] H. S. Leff and A. F. Rex, *Maxwell's Demon 2: Entropy, Classical and Quantum Information, Computing* (Taylor and Francis, 2002).
 - [9] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law, *Int. J. Theor. Phys.* **39**, 2717 (2000).
 - [10] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, R. W. Spekkens, arXiv:1111.3882 (2011).
 - [11] M. Horodecki, J. Oppenheim, (Quantumness in the context of) Resource theories, arXiv:1209.2162 (2012).
 - [12] O. Dahlsten, R. Renner, E. Rieper, and V. Vedral, Inadequacy of von Neumann entropy for characterizing extractable work. *New Journal of Physics*, **13**, 053015 (2011).
 - [13] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy. *Nature* **474**, 61 - 63 (2011).
 - [14] J. Åberg, Truly work-like work extraction, arXiv:1110.6121 (2011).
 - [15] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nano thermodynamics. arXiv:1111.3834 (2011).
 - [16] D. Egloff, O. Dahlsten, R. Renner, and V. Vedral, Laws of thermodynamics beyond the von Neumann regime. arXiv:1207.0434 (2012).
 - [17] D. Egloff, Work value of mixed states in single instance work extraction games, Master's thesis, ETH Zurich (2010).
 - [18] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, A quantitative Landauer's principle, arXiv:1211.1037 (2012).
 - [19] P. Skrzypczyk, A. J. Short, S. Popescu, Extracting work from quantum systems, arXiv:1302.2811 (2013).
 - [20] D. Gelbwaser-Klimovsky, R. Alicki, G. Kurizki, How much work can a quantum device extract from a heat engine? arXiv:1302.3468 (2013).
 - [21] S. D. Bartlett, T. Rudolph, R. W. Spekkens, Reference frames, superselection rules, and quantum information, *Rev. Mod. Phys.* **79**, 555 (2007).
 - [22] G. Gour, R. W. Spekkens, The resource theory of quantum reference frames: manipulations and monotones, *New J. Phys.* **10**, 033023 (2008).
 - [23] I. Marvian, R. Spekkens, The theory of manipulations of pure state asymmetry I: basic tools and equivalence classes of states under symmetric operations, arXiv:1104.0081 (2011).
 - [24] R. J. Glauber, Photon Correlations, *Phys. Rev. Lett.* **10**, 84 (1963).
 - [25] R. J. Glauber, Coherent and incoherent states of the radiation field, *Phys. Rev.* **131**, 2766 (1963).
 - [26] L. Mandel and E. Wolf, *Optical coherence and quantum optics* (Cambridge University Press, Cambridge, 2008).
 - [27] K. Mølmer, Optical coherence: a convenient fiction, *Phys. Rev. A* **55**, 3195 (1997).
 - [28] J. Gea-Banacloche, Comment on "Optical coherence: a convenient fiction", *Phys. Rev. A* **58**, 4244 (1998).
 - [29] K. Mølmer, Reply to "Comment on 'Optical coherence: a convenient fiction' ", *Phys. Rev. A* **58**, 4247 (1998).
 - [30] T. Rudolph, B. C. Sanders, *Phys. Rev. Lett.* **87**, 077903 (2001).
 - [31] S. J. van Enk, C. A. Fuchs, Quantum state of an ideal propagating laser field, *Phys. Rev. Lett.* **88**, 027902 (2002).
 - [32] S. D. Bartlett, T. Rudolph, R. W. Spekkens, Dialogue concerning two views on quantum coherence: factist and fictionist, *Int. J. Quant. Inf.*, **4**, 17 (2006).
 - [33] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).
 - [34] G. Lindblad, *Non-Equilibrium Entropy and Irreversibility* (Reidel, Lancaster, 1983).
 - [35] K. Takara, H.-H. Hasegawa, and D. J. Driebe, Generalization of the second law for a transition between nonequilibrium states, *Phys. Lett. A* **375**, 88 (2010).

- [36] M. Esposito and C. Van den Broeck, Second law and Landauer principle far from equilibrium, *Euro. Phys. Lett.* **95**, 40004 (2011).
- [37] D. Aharonov, A. Kitaev, N. Nisan, Quantum circuits with mixed states, *Proc. 30th ACM Symposium on Theory of Computation*, 20 (1997).
- [38] A. Kitaev, Quantum computations: algorithms and error correction, *Russian Math. Surveys* **52**, 1191 (1997).
- [39] R. A Horn and C. R. Johnson, *Topics in matrix analysis* (Cambridge University Press, Cambridge, 1991).
- [40] J. H. Eberly, N. B. Narozhny, J. J. Sanchez-Mondragon, Periodic spontaneous collapse and revival in a simple quantum model, *Phys. Rev. Lett.* **44**, 1323 (1980).
- [41] R. Alicki, M. Horodecki, P. Horodecki, R. Horodecki, Thermodynamics of quantum information systems – Hamiltonian description, *Open sys. and Information dyn.* **11**, 205 (2004).
- [42] M. J. Henrich, G. Mahler, M. Michel, Driven spin systems as quantum thermodynamic machines: fundamental limits, *Phys. Rev. E* **75**, 051118 (2007).
- [43] O. J. E. Maroney, Generalizing Landauer’s principle, *Phys. Rev. E*, **79**, 031105 (2009),
- [44] Minimal energy cost for thermodynamic information processing: measurement and information erasure, *Phys. Rev. Lett.* **102**, 250602 (2009).
- [45] J. Anders, V. Giovannetti, Thermodynamics of discrete quantum processes, [arXiv:1211.0183](https://arxiv.org/abs/1211.0183) (2012).
- [46] W. Pusz and S. L. Woronowicz, Passive states and KMS states for general quantum systems, *Commun. Math. Phys.* **58**, 273 (1978).
- [47] A. Lenard, Thermodynamical proof of the Gibbs formula for elementary quantum systems, *J. Stat. Phys.* **19**, 575 (1978).
- [48] R. Alicki and M. Fannes, Extractable work from ensembles of quantum batteries. Entanglement helps, [arXiv:1211.1209](https://arxiv.org/abs/1211.1209) (2012).
- [49] K. V. Hovhannisyan, M. Perarnau-Llobet, M. Huber, A. Acín, The role of entanglement in work extraction, [arXiv:1303.4686](https://arxiv.org/abs/1303.4686) (2013).
- [50] R. Bhatia *Matrix Analysis (Graduate texts in Mathematics Vol. 169)* (New York, Springer, 1997).
- [51] R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd Ed. (Cambridge University Press, Cambridge, 2013).
- [52] K. M. R. Audenaert, *J. Phys. A* **40**, 8127 (2007).
- [53] D. Reeb and M. M. Wolf, Tight bound on relative entropy by entropy difference, [arXiv:1304.0036](https://arxiv.org/abs/1304.0036) (2013).
- [54] R. B. Bapat and V. S. Sunder, On majorization and Schur products, *Linear Algebra Appl.* **72**, 107 (1985).
- [55] A. Chefles, Quantum operations, state transformations and probabilities, *Phys. Rev. A* **65**, 052314 (2002).